2000

Perron-Frobenius theory and Perron Complementation

Angela Hang Tran
San Jose State University

PERRON—FROBENIUS THEORY

AND

PERRON COMPLEMENTATION

A Thesis

Presented to

The faculty of the Department of Mathematics

San Jose State University

In Partial Fulfillment

of the Requirements for the Degree

Master of Science

by

Angela Hang Tran

May 2000

UMI Number: 1399826

APPROVED FOR THE DEPARTMENT OF MATHEMATICS

Jane M. Day

Dr. Jane Day

Ho Kuen Ng

Dr. Ho Kuen Ng

Brian Peterson

Dr. Brian Peterson

APPROVED FOR THE UNIVERSITY

William Fish

ABSTRACT

PERRON—FROBENIUS THEORY AND PERRON COMPLEMENTATION

by Angela Hang Tran

This thesis begins with a proof of the Perron-Frobenius theorem, an early study on n x n nonnegative irreducible matrices. Among other things, this theorem says that when A is nonnegative and irreducible, the spectral radius of A, $\rho(A)$, is positive with algebraic multiplicity one, and there is a unique positive eigenvector associated with ρ whose entries sum to one. This vector is called the Perron vector of A.

In a variety of applications, such matrices arise and it is important to calculate the Perron vector. This thesis presents a method developed by C. Meyer for doing this. His method is based on the theory of Perron-Frobenius and M-matrices, and calculates the Perron vector by "uncoupling" the matrix A.

ACKNOWLEDGEMENTS

Thanks to God who has made everything possible for me to complete this thesis in a desirable time.

I am extremely grateful to my advisor, Dr. Jane Day, for giving me a lot of her valuable time to guide me in researching and editing this thesis.

Also, many thanks to Dr. Ho Kuen Ng and Dr. Brian Peterson for being committee members for my thesis, carefully reading the paper, and giving me instructive criticism.

Special thanks to my Dad, who always checks and provides me snacks and drinks while I am studying, especially when I was typing this thesis. Thank to my beloved husband, TuAnh, who inspires me by his love, care, and success. Finally, many thanks to my sister, brothers, teachers, staff, and friends who have helped and supported me during my work on this degree.

CONTENTS

INTRODUCTION

Nonnegative matrices arise in many applications such as discrete economic models, traffic control, scheduling (tasks), etc., so many people are interested in their properties. Early papers by O. Perron and G. Frobenius provided an interesting theory for nonnegative irreducible matrices, now called the Perron-Frobenius theorem. In part one, we will present the background theory needed about irreducibility, graphs, and matrix norms needed for a modern proof of this theorem.

A special type of matrices called M-matrices is closely related to nonnegative matrices. The derivation of the properties of M-matrices rely heavily on the Perron-Frobenius theory. Finally, we will present a recent application of all these ideas due to C. Meyer. He has used the properties of M-matrices together with the Perron-Frobenius theory to create a theory he calls Perron Complementation and a method for using that with parallel processing to determine the dominant eigenvector of a nonnegative irreducible matrix A. (A dominant eigenvector of a matrix is one that corresponds to an eigenvalue which has maximum modulus.) In part two, we will present this theory and new method. Our major sources are the papers [M], [MM], and [MH]. Part two begins with some necessary background material on primitive and M-matrices.

Throughout, M_n will denote the set of all n x n real matrices and $M_{m,n}$ for the set of all m x n real matrices. A permutation matrix is one obtained from the n x n identity matrix by rearranging its rows. It is true that the inverse of a permutation matrix is its transpose.

Let $A = [a_{ij}] \in M_n$ and $x = [x_i] \in R^n$. We will write

A is nonnegative, denoted $A \geq 0$, if all $a_{ij} \geq 0$,

A is positive, denoted $A > 0$, if all $a_{ij} > 0$,

x is positive, denoted $x > 0$, if each $x_i > 0$.

$A \geq B$ or $A > B$ if $A - B \geq 0$ or $A - B > 0$ respectively.

PART ONE

THE PERRON—FROBENIUS THEORY

A. BACKGROUND

I. IRREDUCIBILITY [HJ, chapter 6.2]

(I.A.1.1) **Definition**: A matrix $A \in M_n$ is *reducible* if either $A = 0$ and $n = 1$, or $n \geq 2$ and there is a permutation matrix $P \in M_n$ such that

$PAP^T = \begin{bmatrix} B & C \\ O & D \end{bmatrix}$ where B, D are square matrices of size $r \times r$, $(n-r) \times (n-r)$ respectively, $1 \leq r < n$, and O is the $(n-r) \times r$ zero matrix.

Example 1: $A = \begin{bmatrix} 1 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 \\ 1 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix}$

$$\text{If } P = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}, \text{ then } PAP^T = \left[\begin{array}{cc:ccc} 1 & 1 & 0 & 1 & 1 \\ 1 & 0 & 1 & 1 & 0 \\ \hdashline 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{array}\right],$$

so A is reducible.

(I.A.1.2) **Definition**: A matrix $A \in M_n$ is *irreducible* if it is not reducible.

$$\text{Example 2: } A = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}$$

There are six permutation matrices P of size 3 x 3, and direct checking shows that none of them makes PAP^T have the block upper triangular form required for A to be reducible.

(I.A.1.3) **Notes**:

(a) If each entry of A is nonzero then A is irreducible.

(b) If A is irreducible then A has no zero row or column.

(c) If A is reducible then A has at least (n-1) zero entries.

(I.A.1.4) **Definition**: A matrix $A = [a_{ij}] \in M_n$ has *property SC* if for every pair of distinct integers p, q with $1 \le p,q \le n$ there is a sequence of distinct integers $k_1 = p$, k_2, k_3, ..., k_{m-1}, $k_m = q$, $2 \le m \le n$, such that

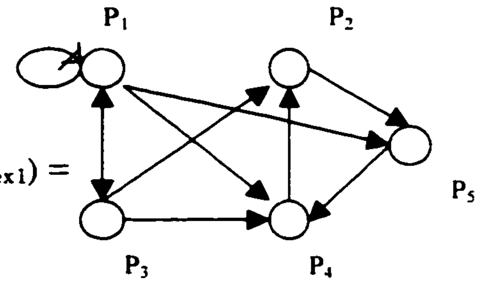
$a_{k_1 k_2} a_{k_2 k_3} \cdots a_{k_{m-1} k_m} \neq 0$.

For example, the matrix in Example 1 does not have property SC because the pair 2,1 does not have such a sequence (notice only paths $a_{25}a_{54}$ and $a_{25}a_{54}a_{42}$ originate at any a_{2j}). However, the matrix in Example 2 does have property SC, which one can check by exhaustion:

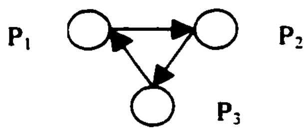
pair 1,2: $a_{12} \neq 0$	pair 2,1: $a_{23}a_{31} \neq 0$	pair 3,1: $a_{31} \neq 0$
pair 1,3: $a_{12}a_{23} \neq 0$	pair 2,3: $a_{23} \neq 0$	pair 3,2: $a_{31}a_{12} \neq 0$

(I.A.1.5) **Definition**: The *directed graph of* $A \in M_n$, denoted by G(A), is the directed graph on n nodes $P_1, P_2, \ldots, P_n$ such that there is a directed arc in G(A) from P_i to P_j if and only if $a_{ij} \neq 0$.

Example: If $A_{ex1} = \begin{bmatrix} 1 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 \\ 1 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix}$ then $G(A_{ex1}) =$



If $A_{ex2} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}$ then $G(A_{ex2}) =$



(I.A.1.6) **Definition**: A *directed path* in a graph G is a sequence of directed arcs $P_{i_1}P_{i_2}$, $P_{i_2}P_{i_3}$, $P_{i_3}P_{i_4}$, ..., each beginning at the end point of the preceding arc. The *ordered list of nodes* in such a directed path is P_{i_1}, P_{i_2}, ... The *length* of a directed path is the number of successive arcs in the directed path if this number is finite; otherwise, the directed path is said to have infinite length. A *cycle* is a directed path that begins and ends at the same node, such that this node occurs exactly twice in the ordered list of nodes in the path, and no other node occurs more than once in the list.

For example, in $G(A_{ex1})$, $P_1P_3P_1$ is a cycle, and in $G(A_{ex2})$, $P_1P_2P_3P_1$ is a cycle.

(I.A.1.7) **Definition**: A directed graph G is *strongly connected* if for every pair of distinct nodes P_i, P_j in G, there is a directed path of finite length that begins at P_i and ends at P_j.

For example, the directed graph $G(A_{ex1})$ is not strongly connected since there is no directed path from node P_2 to P_3. However, the graph $G(A_{ex2})$ is strongly connected.

(I.A.1.8) **Theorem**: A matrix A $\in M_n$ has property SC if and only if the directed graph G(A) is strongly connected.

Proof:

If G(A) is strongly connected, then by definition (I.A.1.7), there is a path of finite length between any nodes P_i and P_j. Deleting any cycle that may appear in this path, we are left with the path whose nodes are distinct starting at P_i and ending at P_j. The indices of the nodes make up a sequence (i, k_2, k_3,.., k_{m-1}, j) of distinct integers, and by the definition of the directed graph of A, we have $a_{ik_1}a_{k_1k_2}a_{k_2k_3}\cdots a_{k_{m-1}j} \neq 0$. So A has property SC. Conversely, if A has property SC, then for every pair of distinct integers p, q with $1 \leq p, q \leq n$ there is a sequence of distinct integers $k_1 = p, k_2, k_3, \ldots, k_{m-1}, k_m = q$, $1 \leq m \leq n$, such that $a_{k_1k_2}a_{k_2k_3}\cdots a_{k_{m-1}k_m} \neq 0$. This implies each entry $a_{k_rk_{r-1}} \neq 0$; so, there is a path from P_p to P_q for any $p \neq q$. Thus, G(A) is strongly connected.

□

(I.A.1.9) **Lemma**: If A $\in M_n$ and m >1, the (i,j) entry of A^m equals

$$\sum_{1\leq k_1,\ldots,k_{m-1}\leq n} a_{ik_1}a_{k_1k_2}a_{k_2k_3}\cdots a_{k_{m-1}j} .$$

Proof:

We will induct on m. When m = 2, $(A^2)_{ij} = \sum_{1\leq s\leq n} a_{is}a_{sj}$ by the definition of matrix product. Let m ≥ 2 and assume $\left(A^m\right)_{ij} = \sum_{1\le k_1,\dots,k_{m-1}\le n} a_{ik_1}a_{k_1k_2}a_{k_2k_3}\cdots a_{k_{m-1}j}$.

$$\begin{aligned}\text{Then } (A^{m+1})_{ij} &= \sum_{1\le t\le n}\left(A^m\right)_{it}a_{tj} \\ &= \sum_{1\le t\le n}\left(\sum_{1\le k_1,\dots,k_{m-1}\le n} a_{ik_1}a_{k_1k_2}a_{k_2k_3}\cdots a_{k_{m-1}t}\right)a_{tj} \\ &= \sum_{1\le k_1,\dots,k_{m-1},t\le n} a_{ik_1}a_{k_1k_2}a_{k_2k_3}\cdots a_{k_{m-1}t}a_{tj} .\end{aligned}$$

Rename $t = k_m$; then $(A^{m+1})_{ij} = \sum_{1\le k_1,\dots,k_{m-1},k_m\le n} a_{ik_1}a_{k_1k_2}a_{k_2k_3}\cdots a_{k_{m-1}k_m}a_{k_mj}$.

□

(I.A.1.10) **Definition**: When A is a matrix, $|A|$ will denote $[\,|a_{ij}|\,]$.

(I.A.1.11) **Theorem**: Let $A \in M_n$ and let P_i and P_j be given nodes in G(A).

There exists a directed path of length m in G(A) from P_i to P_j if and only if $(|A|^m)_{ij} \neq 0$.

Proof:

By lemma (A.1.9), $(|A|^m)_{ij} = \sum_{1\le k_1,\dots,k_{m-1}\le n} |a_{ik_1}||a_{k_1k_2}||a_{k_2k_3}|\cdots|a_{k_{m-1}j}|$. Thus,

$(|A|^m)_{ij} \neq 0$ if and only if $a_{ik_1}a_{k_1k_2}a_{k_2k_3}\cdots a_{k_{m-1}j} \neq 0$ for some $1 \le k_1, k_2, .., k_{m-1} \le n$.

This is equivalent to having a path from P_i to P_j of length m.

□

(I.A.1.12) **Corollary**: If $A \in M_n$, then G(A) is strongly connected if and only if

$(I + |A|)^{n-1} > 0$.

Proof:

Observe that $(I+|A|)^{n-1} = I + \binom{n-1}{1}|A| + \binom{n-1}{2}|A|^2 + \ldots + \binom{n-1}{n-1}|A|^{n-1}$. This is positive if and only if for each pair (i,j) with $i \neq j$, at least one of the terms $|A|$, $|A|^2$, ..., $|A|^{n-1}$ has a positive (i,j) entry. By Theorem (I.A.1.11), this is equivalent to the existence of a directed path in G(A) from P_i to P_j, i.e., to G(A) being strongly connected.

□

(I.A.1.13) **Theorem**: If $A \in M_n$ and $n > 1$, then A is irreducible if and only if

$(I + |A|)^{n-1} > 0$.

Proof:

If n = 1, this is clear. Suppose now that $n \geq 2$. We will prove the contrapositive; that is, A is reducible if and only if $(I + |A|)^{n-1}$ has at least one zero entry. Suppose A is reducible, then there is a permutation matrix $P \in M_n$ such that $\tilde{A} = PAP^T = \begin{bmatrix} B & C \\ O & D \end{bmatrix}$ where B, C, O, and D are block matrices as in the Definition (I.A.1.1). Since P and P^T simply permute rows and columns, $|A| = |P^T \tilde{A} P| = P^T |\tilde{A}| P$. Since $\tilde{A}$ has an (n–r) x r (with $r \geq 1$) zero block in the lower left corner, all $|\tilde{A}|^2, |\tilde{A}|^3, \ldots, |\tilde{A}|^{n-1}$ have the same zero block in the lower left corner.

Thus, $(I+|A|)^{n-1} = (I+P^T|\tilde{A}|P)^{n-1} = P^T(I+|\tilde{A}|)^{n-1}P$

$$= P^T\left[I+(n-1)|\tilde{A}|+\binom{n-1}{2}|\tilde{A}|^2+\ldots+\binom{n-1}{n-1}|\tilde{A}|^{n-1}\right]P$$

and all the terms in the square brackets have an (n-r) x r block of zero's in the lower left corner (with $r \geq 1$), so, $(I+|A|)^{n-1}$ has at least one zero entry.

Conversely, suppose the (i,j) entry of $(I+|A|)^{n-1}$ is zero; then $i \neq j$ since the diagonal entries are positive from the term I in the sum. By Theorem (I.A.1.11), there is no directed path in $G(I+|A|)$ or in fact, in G(A) (since $i \neq j$), from node P_i to node P_j. Define the set of nodes $S_1 = \{P_p \mid P_p = P_j$ or there is a path in G(A) from P_p to $P_j\}$. Then S_1 is not empty because P_j is an element of S_1. Let S_2 be the set of all nodes in G(A) that are not in S_1. Then, S_2 is not empty either since P_i is not an element of S_1.

Let r be the number of elements in S_1. Then S_2 has $n-r > 0$ elements. Also, there is no path from any node P_q in S_2 to P_p in S_1. For if there were, then (by definition of S_1) there would be a path $P_q \rightsquigarrow P_p \rightsquigarrow P_j$, so P_q would already be in S_1, a contradiction.

Let $e_1, e_2, \ldots, e_n$ denote the rows of I_n, and form a permutation matrix P as follows:

a. the first r rows are $e_{p_1},\cdots,e_{p_r}$ where $S_1 = \{P_{p_1},\cdots,P_{p_r}\}$.

b. the last (n-r) rows are $e_{q_1},\cdots,e_{q_{n-r}}$ where $S_2 = \{Q_{q_1},\cdots,Q_{q_{n-r}}\}$.

(See example after this proof.)

Then, PAP^T will have

a. all columns and rows of A corresponding to the index of a node in S_1 in the upper left corner, forming B (size r x r);

b. all columns and rows of A corresponding to the index of a node in S_2 in the lower right corner, forming D (size (n-r) x (n-r));

c. a matrix C of size r x (n-r) in the upper right corner;

d. a zero block, O, of size (n-r) x r in the lower left corner.

That is $PAP^T = \begin{bmatrix} B & C \\ O & D \end{bmatrix}$, so A is reducible.

□

(I.A.1.14) **Example**:

Consider the graph $G(A_{ex1})$ of the example in (I.A.1.5). We see that if (i,j) = (3,1), only nodes P_1 and P_3 have a path to P_1. So $S_1 = \{P_1, P_3\}$ and $S_2 = \{P_2, P_4, P_5\}$.

Then, by the above algorithm, $P = \begin{bmatrix} e_1 \\ e_3 \\ e_2 \\ e_4 \\ e_5 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$

This is exactly the same as the matrix P in Example 1.

(I.A.1.15) **Remark**:

It is very hard to check directly whether or not a large graph is strongly connected. The above theorem provides a preferred tool since matrix multiplication can be done quickly by software like Matlab. For example, let

$$A = \begin{bmatrix} 0 & 0 & 0 & 2 & 0 & 0 & 1 & -1 \\ 0 & 0 & 1 & 0 & 0 & 0 & -2 & 3 \\ 1 & -1 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & -2 \\ 0 & 0 & 0 & 0 & 0 & 0 & -3 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 4 & 0 & 4 & -3 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

Then A is irreducible and its graph is strongly connected since $(I+|A|)^7$ is the following matrix:

$$(I+|A|)^7 = \begin{bmatrix} 1726 & 1381 & 7099 & 12990 & 9517 & 1725 & 13984 & 16783 \\ 3522 & 29987 & 15141 & 28126 & 20679 & 3522 & 30263 & 36097 \\ 1412 & 12560 & 6136 & 11826 & 8700 & 1412 & 11724 & 14000 \\ 202 & 1306 & 678 & 1289 & 920 & 202 & 1224 & 1518 \\ 4320 & 37860 & 19068 & 35316 & 26044 & 4320 & 38460 & 45708 \\ 888 & 9204 & 4320 & 8604 & 6388 & 889 & 8385 & 9864 \\ 6356 & 57636 & 27856 & 54160 & 39900 & 6356 & 52754 & 62844 \\ 339 & 2787 & 1412 & 2650 & 1937 & 339 & 2787 & 3349 \end{bmatrix}.$$

We summarize (I.A.1.8), (I.A.1.11), and (I.A.1.13) in the following theorem:

(I.A.1.16) **Theorem**: A matrix $A \in M_n$, where $n > 1$, is irreducible

if and only if

a. $(I + |A|)^{n-1} > 0$, or

b. $G(A)$ is strongly connected, or

c. A has property SC.

II. MATRIX NORMS [HJ, chapter 5.6]

(I.A.2.1) **Definition**: Call a function $\|\bullet\|: M_n \to R$ a *matrix norm* if for all A, B $\in M_n$ it satisfies the following five axioms:

(a) $\|A\| \geq 0$ (Nonnegative)

(b) $\|A\| = 0$ if and only if A = O (Positive)

(c) $\|cA\| = |c|\|A\|$ for all complex scalars c (Homogeneous)

(d) $\|A+B\| \leq \|A\| + \|B\|$ (Triangular inequality)

(e) $\|AB\| \leq \|A\| \, \|B\|$ (Submultiplicative)

Notice (a)—(d) are the usual conditions for a vector norm.

(I.A.2.2) **Definition**:

(a) *The maximum column sum matrix norm* $\|\bullet\|_1$ is defined on M_n by

$$\|A\|_1 = \max_{1 \leq j \leq n} \sum_{i=1}^{n} |a_{ij}|.$$

(b) *The maximum row sum matrix norm* $\|\bullet\|_\infty$ is defined on M_n by

$$\|A\|_\infty = \max_{1 \leq i \leq n} \sum_{j=1}^{n} |a_{ij}|.$$

(c) The *Euclidean norm* $\|\bullet\|_2$ or *l_2 norm* is defined for A $\in M_n$ by

$\|A\|_2 = \left(\sum_{i,j=1}^{n} |a_{ij}|^2 \right)^{1/2}$. (Some authors call this the Frobenius norm, and use the notation $\|A\|_2$ for a different norm.)

For example, if A = $\begin{bmatrix} 1 & 0 & 2 \\ -2 & 1 & 0 \\ 0 & -1 & 0 \end{bmatrix}$, then $\|A\|_1 = 4$, $\|A\|_\infty = 3$, and $\|A\|_2 = \sqrt{11}$.

(I.A.2.3) **Definition**: The *spectral radius* of A ∈ M_n, denoted ρ(A), is the largest of the moduli of the eigenvalues of A.

For example, if the eigenvalues of a matrix A ∈ M_4 are −2, 1, and 1 ± i, then their moduli are 2, 1, and $\sqrt{2}$, so ρ(A) = 2.

(I.A.2.4) **Theorem**: If $\|\bullet\|$ is any matrix norm and if A ∈ M_n, then ρ(A) ≤ $\|A\|$.

Proof:

If Ax = λx, x ≠ 0 and | λ | = ρ(A), let X ∈ M_n be the matrix each of whose columns is equal to the vector x. Observe that AX = λX. If $\|\bullet\|$ is any matrix norm, then $|\lambda|\,\|X\| = \|\lambda X\| = \|AX\| \le \|A\|\,\|X\|$. Since x ≠ 0, $\|X\| > 0$; therefore, ρ(A) = | λ | ≤ $\|A\|$.

□

(I.A.2.5) **Lemma**: Let $A \in M_n$ be a nonnegative matrix. If the row sums of A are constant then $\rho(A) = \|A\|_\infty$. If the column sums of A are constant then $\rho(A) = \|A\|_1$.

Proof:

If the row sums are constant, then $x = [1 \ldots 1]^T$ is an eigenvector with eigenvalue $\|A\|_\infty$. But we know that $\rho(A) \le \|A\|_\infty$ by Theorem (I.A.2.4); thus, $\rho(A) = \|A\|_\infty$. Apply this to A^T to obtain the statement for column sums.

□

(I.A.2.6) **Definition**: A matrix A is called *convergent* if $A^m \to 0$ as $m \to \infty$.

(I.A.2.7) **Theorem**: A matrix A is convergent if and only if $\rho(A) < 1$.

Proof:

We will use the Jordan canonical form theorem [HJ, p.126]. That is, there is a nonsingular matrix S such that $SAS^{-1} = \begin{bmatrix} J_1 & 0 & \cdots & & 0 \\ 0 & J_2 & 0 & \cdots & 0 \\ \vdots & 0 & \ddots & \ddots & \vdots \\ & \vdots & \ddots & & 0 \\ 0 & 0 & \cdots & 0 & J_k \end{bmatrix}$

where $J_i = \begin{bmatrix} \lambda_i & 1 & 0 & \cdots & & 0 \\ 0 & \lambda_i & 1 & 0 & \cdots & 0 \\ & 0 & \ddots & \ddots & & \\ \vdots & \vdots & \ddots & \ddots & \ddots & 0 \\ & & & 0 & \lambda_i & 1 \\ 0 & 0 & \cdots & & 0 & \lambda_i \end{bmatrix} = \lambda_i I + N$, and $N = \begin{bmatrix} 0 & 1 & 0 & \cdots & & 0 \\ 0 & 0 & 1 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & & \vdots \\ & & & & & 0 \\ & & & & 0 & 1 \\ 0 & 0 & \cdots & & & 0 \end{bmatrix}$.

Suppose N is r x r. Then $N^r = 0$, hence if $m \geq r$, we have

$J_i^m = \lambda_i^m I + \binom{m}{1}\lambda_i^{m-1}N + \cdots + \binom{m}{r-1}\lambda_i^{m-r+1}N^{r-1}$. Therefore $J_i^m \to 0$ if and only if

$|\lambda_i| < 1$. Finally, observe that $A^m \to 0$ if and only if $SA^mS^{-1} \to 0$, if and only if $J_i^m \to 0$ for each i, hence $A^m \to 0$ if and only if $|\lambda_i| < 1$ for each eigenvalue λ_i of A.

□

(I.A.2.8) **Lemma**: Let $A \in M_n$ and let $\|\bullet\|$ be any matrix norm.

Then $\rho(A) = \lim_{k\to\infty} \|A^k\|^{1/k}$.

Proof:

Using the Jordan canonical form theorem for A, we see that the eigenvalues of A^k are all λ^k where λ is an eigenvalue of A. Therefore, $\rho(A)^k = \rho(A^k)$. By Theorem (I.A.2.4), $\rho(A^k) \leq \|A^k\|$, so we have $\rho(A) \leq \|A^k\|^{1/k}$ for all k = 1, 2, ... If $\varepsilon > 0$ is given, let $B = [\rho(A)+ \varepsilon]^{-1}A$. Then $\rho(B) = [\rho(A)+ \varepsilon]^{-1}\rho(A) < 1$, hence B is convergent by Lemma (I.A.2.7), i.e. $B^k \to O$ as $k \to \infty$. Thus, $\|B^k\| \to \|O\| = 0$ as $k \to \infty$, hence there is some large integer N so that for all k ≥ N, $\|B^k\| < 1$. Also,

$\|B^k\| = \| [\rho(A)+\varepsilon]^{-k} A^k \| = [\rho(A)+\varepsilon]^{-k}\|A^k\|$; therefore $\|A^k\| \leq [\rho(A)+\varepsilon]^k$ for all

k ≥ N, hence $\|A^k\|^{1/k} \leq \rho(A)+\varepsilon$ for all k ≥ N. Since $\rho(A) \leq \|A^k\|^{1/k}$ for all

k = 1, 2, ... and ε > 0 is arbitrary, we can conclude that $\lim\limits_{k\to\infty} \|A^k\|^{1/k}$ exists and

equals ρ(A).

□

(I.A.2.9) **Lemma**: Let A, B ∈ $M_{m,n}$ and 0 ≤ A ≤ B. Then $\|A\|_2 \leq \|B\|_2$.

Proof:

Since $0 \leq a_{ij} \leq b_{ij}$ for all $1 \leq i \leq m$ and $1 \leq j \leq n$,

$$\sum_{\substack{1\leq i\leq m\\ 1\leq j\leq n}} | a_{ij} |^2 \leq \sum_{\substack{1\leq i\leq m\\ 1\leq j\leq n}} | b_{ij} |^2 \Leftrightarrow \left(\sum_{\substack{1\leq i\leq m\\ 1\leq j\leq n}} | a_{ij} |^2 \right)^{1/2} \leq \left(\sum_{\substack{1\leq i\leq m\\ 1\leq j\leq n}} | b_{ij} |^2 \right)^{1/2} \Leftrightarrow \|A\|_2 \leq \|B\|_2 .$$

□

B. THE PERRON—FROBENIUS THEORY

In 1907, O. Perron published fundamental results for positive matrices (every entry is positive). In 1912, G. Frobenius extended these results to nonnegative irreducible matrices [BP, p.27]. His extension, usually called the Perron-Frobenius Theorem, is less well known. We will supply a proof and show how the classic Perron theorem follows from this more general theorem. The proof makes use of some simple graph theory and is based on the development in [HJ chapter 8, sections 2, 3, and 4].

(I.B.1) **Theorem**: If A, B $\in M_n$ are nonnegative and $A \leq B$, then $\rho(A) \leq \rho(B)$.

Proof:

Since $0 \leq A \leq B$, we have $A^m \leq B^m$ for m = 1, ..., hence $\|A^m\|_2 \leq \|B^m\|_2$ by Lemma (I.A.2.9). Taking roots and then letting $m \to \infty$, we see that $\rho(A) \leq \rho(B)$ by Theorem (I.A.2.8).

□

(I.B.2) **Theorem**: If A $\in M_n$ is nonnegative then

$$\min_{1\le i\le n}\sum_{j=1}^{n} a_{ij} \le \rho(A) \le \max_{1\le i\le n}\sum_{j=1}^{n} a_{ij} \text{ and } \min_{1\le j\le n}\sum_{i=1}^{n} a_{ij} \le \rho(A) \le \max_{1\le j\le n}\sum_{i=1}^{n} a_{ij}.$$

Proof:

For each row i, let $s_i = \sum_{j=1}^{n} a_{ij}$ and $s = \min_{1\le i\le n}\{s_i\}$. To see the minimum row sum is a lower bound for ρ(A), construct a matrix B such that $0 \le B \le A$ and $\sum_{j=1}^{n} b_{ij} = s$ for all i = 1, ..., n as follows: B = O if s = 0, else $B = [b_{ij}] = \left[\frac{s}{s_i} a_{ij}\right]$. This construction of B gives ρ(B) = s by Lemma (I.A.2.5), and ρ(B) ≤ ρ(A) by Theorem (I.B.1).

Similarly, if t is the maximum row sum of A, construct a matrix B such that $0 \le A \le B$ and $\sum_{j=1}^{n} b_{ij} = t$ for all i = 1, ..., n. Then t = ρ(B), so the upper bound is established.

The column sum bounds follow from applying the row sum bounds to A^T.

□

(I.B.3) **Theorem**: Let $A \in M_n$ be an irreducible nonnegative matrix. Then ρ(A) > 0. In particular, if A > 0 then ρ(A) > 0.

Proof:

The minimum row sum of an irreducible nonnegative matrix must be positive by (I.A.1.3). Thus, ρ(A) > 0 by Theorem (I.B.2).

□

(I.B.4) **Corollary**: Let $A \in M_n$ be nonnegative and $x \in R^n$ be positive. Then

$$\min_{1\le i\le n} \frac{1}{x_i}\sum_{j=1}^{n} a_{ij}x_j \le \rho(A) \le \max_{1\le i\le n} \frac{1}{x_i}\sum_{j=1}^{n} a_{ij}x_j \quad \text{and}$$

$$\min_{1\le j\le n} x_j\sum_{i=1}^{n} (a_{ij}/x_i) \le \rho(A) \le \max_{1\le j\le n} x_j\sum_{i=1}^{n} (a_{ij}/x_i).$$

Proof:

Let $S = \begin{bmatrix} x_1 & & 0 \\ & \ddots & \\ 0 & & x_n \end{bmatrix}$. Since all $x_i > 0$, S is nonsingular. By similarity, $\rho(S^{-1}AS) = \rho(A)$. Also, $S^{-1}AS$ is nonnegative since S, S^{-1}, and A are. Applying (I.B.2) to $S^{-1}AS$, we get the desired inequalities.

□

(I.B.5) **Corollary**: Let $A \in M_n$ be nonnegative and $x \in C^n$ be positive.

If $r, s \ge 0$ such that $rx \le Ax \le sx$, then $r \le \rho(A) \le s$. Further, if $rx < Ax$, then $r < \rho(A)$ and $\rho(A) < s$ if $Ax < sx$.

Proof:

If $rx \le Ax \le sx$ then $r \le \min_{1\le i\le n} \frac{1}{x_i}\sum_{j=1}^{n} a_{ij}x_j$ and $\max_{1\le i\le n} \frac{1}{x_i}\sum_{j=1}^{n} a_{ij}x_j \le s$. So by Corollary (I.B.4), we can conclude that $r \le \rho(A) \le s$.

If $rx < Ax$, then there is some $r' > r$ such that $r'x \le Ax$. From previous result, $r' \le \rho(A)$, hence, $r < \rho(A)$. The upper bound is verified similarly.

□

(I.B.6) **Corollary**: Let $A \in M_n$ be nonnegative. If A has a positive eigenvector, then the corresponding eigenvalue is $\rho(A)$; that is, if $A \geq 0$ and $Ax = \lambda x$ where $x > 0$, then $\lambda = \rho(A)$.

Proof:

Since $A \geq 0$ and $x > 0$, we have $Ax \geq 0$. Since we also know $Ax = \lambda x$, λ is real and $\lambda \geq 0$. But $\lambda x \leq Ax \leq \lambda x$, so by Corollary (I.B.5), $\lambda \leq \rho(A) \leq \lambda$, or $\lambda = \rho(A)$.

□

(I.B.7) **Lemma**: Let $A \in M_n$ and $x \in R^n$. If $A > 0$, $x \geq 0$, and $x \neq 0$, then $Ax > 0$.

Proof:

Each entry of Ax is positive because $(Ax)_i = \sum_{k=1}^{n} a_{ik} x_k$, all $a_{ik} > 0$, $x_k \geq 0$, and some $x_k \neq 0$.

□

(I.B.8) **Lemma**: Suppose that $A \in M_n$ is a positive matrix, $0 \neq x \in C^n$, and $Ax = \lambda x$ where $|\lambda| = \rho(A)$. Then $A|x| = \rho(A)|x|$ and the vector $|x|$ is positive.

Proof:

By Theorem (I.B.3), $\rho(A)$ is positive. By Lemma (I.B.7) applied to A and $0 \neq |x| \geq 0$, $A|x| > 0$.

From the hypothesis, we see that $|Ax| = |\lambda x| = |\lambda|\,|x| = \rho(A)|x|$. It is easy to verify that $|Ax| \le A|x|$, so if we let $y = A|x| - \rho(A)|x|$, then $y \ge 0$. If $y_i > 0$ for some i, we will show that this leads to a contradiction. Let $z = A|x|$, so $z > 0$. Because $y \ne 0$, $Ay > 0$ by (I.B.7), and $Ay = Az - \rho(A)z$, so $Az - \rho(A)z > 0$. But $z > 0$ and $Az > \rho(A)z$ imply $\rho(A) > \rho(A)$ by Corollary (I.B.5), which is a contradiction. Thus, $y = A|x| - \rho(A)|x|$ must be the zero vector, i.e. $A|x| = \rho(A)\,|x|$ and therefore $|x|$ is positive.

□

(I.B.9) **Theorem**: If $A \in M_n$ is nonnegative then $\rho(A)$ is an eigenvalue of A and there is a nonnegative vector x, $x \ne 0$, such that $Ax = \rho(A)x$.

Proof:

Let $\varepsilon_n = 1/n$ and $A_{\varepsilon_k} = [a_{ij} + \varepsilon_k]$. Then $A_{\varepsilon_k} > 0$ and $A_{\varepsilon_k} \to A$ as $k \to \infty$. Apply (I.B.8) to A_{ε_k} to get a positive eigenvector x^{ε_k} corresponding to $\rho(A_{\varepsilon_k})$. We may assume $\|x^{\varepsilon_k}\| = 1$ for each k. The set $\{x^{\varepsilon_k}$; k = 1, 2, ...} is a subset of the compact set $\{x \mid x \in C^n$ and $\|x\| = 1\}$. Thus, there is a bounded monotone decreasing subsequence $(\varepsilon'_1, \varepsilon'_2, \ldots)$ of $(\varepsilon_1, \varepsilon_2, \ldots)$ with $\varepsilon'_k \to 0$, such that $x^{\varepsilon'_k}$ converges to some vector x in the compact set. Since each $x^{\varepsilon'_k} > 0$, we have $x \ge 0$, and $x \ne 0$ because $\|x\| = 1$. Further, the bounded monotone decreasing sequence $\{\varepsilon'_k\}_{k=1,2,\ldots}$ creates a bounded monotone decreasing sequence $\{A_{\varepsilon'_k}\}_{k=1,2,\ldots}$, and hence, the sequence of real numbers $\{\rho(A_{\varepsilon'_k})\}_{k=1,2,\ldots}$ is also monotone decreasing and bounded below by $\rho(A)$, by (I.B.1). Thus, $\rho = \lim_{k\to\infty} \rho(A_{\varepsilon'_k})$ exists and $\rho \geq \rho(A)$. Because matrixmultiplication is continuous, we have

$\lim_{k\to\infty} [A_{\varepsilon'_k} x^{\varepsilon'_k}] = \lim_{k\to\infty} A_{\varepsilon'_k} \lim_{k\to\infty} x^{\varepsilon'_k} = Ax.$

We also have $\lim_{k\to\infty} [A_{\varepsilon'_k} x^{\varepsilon'_k}] = \lim_{k\to\infty} [\rho(A_{\varepsilon'_k}) x^{\varepsilon'_k}] = \rho x.$

Therefore $Ax = \rho x$. Since $x \neq 0$ and $\rho \geq \rho(A)$, x is an eigenvector and ρ must equal $\rho(A)$.

□

(I.B.10) **Theorem:** Let $A \in M_n$, $\lambda \in C$, and $x, y \in C^n$, written as columns, such that $Ax = \lambda x$, $A^T y = \lambda y$, and $x^T y = 1$. Define $L := xy^T$. Then

a. $Lx = x$ and $y^T L = y^T$;

b. $L^m = L$ for $m = 1,2, \ldots$;

c. $A^m L = LA^m = \lambda^m L$ for $m = 1,2, \ldots$;

d. $L(A - \lambda L) = 0$;

e. $(A - \lambda L)^m = A^m - \lambda^m L$ for $m = 1,2, \ldots$;

f. Every nonzero eigenvalue of $A - \lambda L$ is an eigenvalue of A and the same vector is an associated eigenvector for both;

g. If λ is an eigenvalue of A such that $|\lambda| = \rho(A) > 0$ and λ has geometric multiplicity one, then λ is not an eigenvalue of $A - \lambda L$;

Suppose now that λ is the only eigenvalue of A with modulus $\rho(A)$ and that the geometric multiplicity of λ is one. Then we also have

h. $\rho(A - \lambda L) < \rho(A)$;

i. $(A/\lambda)^m \rightarrow L$ as $m \rightarrow \infty$.

Proof:

a. $Lx = xy^Tx = x(1) = x$ and $y^TL = y^Txy^T = (1)y^T = y^T$.

b. $L^m = (xy^T)^m = xy^Txy^T \ldots xy^T = x(1)^{m-1}y^T = xy^T = L$.

c. $A^mL = A^mxy^T = A^{m-1}(\lambda x)y^T = \lambda(A^{m-1}x)y^T = \ldots = \lambda^m xy^T = \lambda^m L$. Similarly, $LA^m = xy^TA^m = x(\lambda y^T)A^{m-1} = \lambda x(y^TA^{m-1}) = \ldots = \lambda^m xy^T = \lambda^m L$.

d. $L(A - \lambda L) = LA - \lambda L^2 = \lambda L - \lambda L = 0$ by (c) and (b).

e. We induct on m. When $m = 1$, $(A - \lambda L)^1 = A^1 - \lambda^1 L$. Assume that $(A - \lambda L)^m = A^m - \lambda^m L$. Then,

$$
\begin{aligned}
(A - \lambda L)^{m+1} &= (A - \lambda L)^m (A - \lambda L) = (A^m - \lambda^m L)(A - \lambda L) \\
&= A^{m+1} - \lambda A^m L - \lambda^m L(A - \lambda L) \\
&= A^{m+1} - \lambda A^m L, \text{ by (d)} \\
&= A^{m+1} - \lambda(\lambda^m L), \text{ by (c)} \\
&= A^{m+1} - \lambda^{m+1} L.
\end{aligned}
$$

f. Assume $(A - \lambda L)w = \eta w$, where $\eta \neq 0$ and $w \neq 0$ (zero vector). By (d), $0w = L(A - \lambda L)w = L\eta w$. Thus, $\eta Lw = 0$; and since $\eta \neq 0$, $Lw = 0$. Thus, $\eta w = (A - \lambda L)w = Aw - \lambda Lw = Aw - 0 = Aw$. So, η is also an eigenvalue of A with the same eigenvector.

g. We prove by contradiction. Suppose λ is an eigenvalue of $(A - \lambda L)$ and $(A - \lambda L)w = \lambda w$, $w \neq 0$. Since $\lambda \neq 0$, then by (f), $Aw = \lambda w$. But $Ax = \lambda x$ and the geometric multiplicity of λ is one, so $w = \delta x$ for some $\delta \neq 0$. Therefore, $Aw = (A - \lambda L)w \Leftrightarrow \lambda w = \lambda w - \lambda Lw \Leftrightarrow \lambda Lw = 0 \Leftrightarrow \lambda xy^T\delta x = 0 \Leftrightarrow \lambda\delta x = 0$. This is a contradiction because $\lambda\delta \neq 0$ ($\lambda \neq 0$ because $|\lambda| = \rho(A) > 0$) and $x \neq 0$.

h. If $\rho(A - \lambda L) = 0$, then $\rho(A - \lambda L) < \rho(A)$ since $\rho(A) > 0$. If $\rho(A - \lambda L) > 0$, then by (f), $\rho(A - \lambda L) = |\mu|$ for some nonzero eigenvalue μ of A. If $|\mu| = \rho(A)$ then $\mu = \lambda$ (because λ is the only dominant eigenvalue), but this cannot happen by (g). So, $|\mu| < |\lambda|$, hence (h) is established.

i. From (h), $\dfrac{\rho(A-\lambda L)}{\rho(A)} < 1$. So,

$$\rho\left(\frac{A}{\lambda} - L\right) = \rho\left(\frac{A-\lambda L}{\lambda}\right) = \frac{\rho(A-\lambda L)}{|\lambda|} = \frac{\rho(A-\lambda L)}{\rho(A)} < 1.$$ Then, by Theorem (I.A.2.7), $(A/\lambda - L)^m \to 0$ as $m \to \infty$. Now, apply (e) to get (i).

□

(I.B.11) **Lemma**: Let $z_1, \ldots, z_n$ are complex numbers and $c_1, \ldots, c_n$ are positive numbers. If $\left|\sum_{i=1}^{n} c_i z_i\right| = \sum_{i=1}^{n} c_i |z_i|$ then $|z_i| = z_i$ for all $i = 1, \ldots, n$.

Proof:

Since $\left|\sum_{i=1}^{n} c_i z_i\right| = \sum_{i=1}^{n} c_i |z_i|$, the triangle inequality shows that all vectors $c_i z_i$ must lie on the same ray in the complex plane. Because $\sum_{i=1}^{n} c_i |z_i|$ is real and positive, this ray must be the positive real axis. That is $z_i \geq 0$ for each i, or $|z_i| = z_i$.

□

(I.B.12) **Definition** [HJ, p. 12]: If $A \in M_{m,n}$ then the *rank of A*, denoted rank(A), is the largest number of columns of A that constitute a linearly independent set.

Remark: It is true that rank(A) = rank(A^T), the largest number of rows of A which constitute a linearly independent set.

(I.B.13) **Theorem**: If $A > 0$ and λ is any eigenvalue of A other than $\rho(A)$, then $|\lambda| < \rho(A)$ and $\rho(A)$ has algebraic multiplicity one. Furthermore, the same conclusions hold if $A \geq 0$ but $A^m > 0$ for some $m > 1$.

Proof:

Suppose $A > 0$, $Ax = \lambda x$, $x \neq 0$, and $|\lambda| = \rho(A)$. By (I.B.8), we also have $A|x| = \rho(A)|x|$. Together, these facts imply, for any component of x:

$\rho(A)\,|x_k| = |\lambda|\,|x_k| = |\lambda x_k| = \left|\sum_{p=1}^{n} a_{kp}\, x_p\right| = \sum_{p=1}^{n} a_{kp}\,|x_p|$. By Lemma (I.B.11), each $|x_k| = x_k$. So, $\lambda x = Ax = A|x| = \rho(A)|x| = \rho(A)x$; hence, $\lambda = \rho(A)$ and $\rho(A)$ has geometric multiplicity one.

Now, let s be the algebraic multiplicity of $\rho(A)$. Also, let J be the Jordan Normal Form (JNF) of $\frac{1}{\rho(A)}A$, so $J = S^{-1}\left(\frac{1}{\rho(A)}A\right)S$ for some nonsingular S,

and we may choose S so that $J = \begin{bmatrix} 1 & & & & & * \\ & \ddots & & & & \\ & & 1 & & & \\ & & & \lambda_{s+1} & & \\ & & & & \ddots & \\ 0 & & & & & \lambda_n \end{bmatrix}$, where $|\lambda_i| \le 1$ for all $i > s$. By (I.B.10), $\left(\frac{1}{\rho(A)}A\right)^m$ converges to a rank one matrix, which implies that J^m also converges and the limit has rank one. The form of J clearly implies that the limit of J^m has rank at least s. Therefore s must be one.

Now, suppose $A \ge 0$ and $A^m > 0$ for some m. Using the JNF of A, we see that the eigenvalues of A^m are those of A raised to the power m. Thus, $\rho(A^m) = \rho(A)^m$ and if $\rho(A)$ has algebraic multiplicity s for A, then $\rho(A)^m$ has algebraic multiplicity at least s for A^m. Therefore, s must be one because $A^m > 0$.

□

(I.B.14) **Lemma**: Let $\lambda_1, \ldots, \lambda_n$ be the eigenvalues of $A \in M_n$ including any multiplicities. Then $1 + \lambda_1, \ldots, 1 + \lambda_n$ are the eigenvalues of $I + A$, hence $\rho(I + A) \leq 1 + \rho(A)$. If $A \geq 0$, then $\rho(I + A) = 1 + \rho(A)$.

Proof:

Let λ be an eigenvalue of A having multiplicity k. Then, λ is a root of multiplicity k of the characteristic polynomial $p_A(t) = \det(tI - A) = \det[(t+1)I - (I+A)]$. From the Jordan Normal Form of A, we see that $1 + \lambda_1, \ldots, 1 + \lambda_n$ are the eigenvalues of $I + A$, with the same multiplicities as $\lambda_1, \ldots, \lambda_n$ respectively. Also, we have $\rho(I + A) = \max |1 + \lambda_i| \leq 1 + \max |\lambda_i| \leq 1 + \rho(A)$.

If $A \geq 0$ then $\rho(A)$ is an eigenvalue of A by Theorem (I.B.9) and there is a nonnegative vector x, $x \neq 0$, such that $Ax = \rho(A)x$. Thus, $(I + A)x = [1 + \rho(A)]x$; hence, $1 + \rho(A)$ is an eigenvalue of $(I + A)$; thus, $\rho(I + A) = 1 + \rho(A)$.

□

(I.B.15) **Theorem**: If $A \in M_n$ is irreducible and nonnegative then $\rho(A)$ is an eigenvalue of A and there is a positive vector x such that $Ax = \rho(A)x$. In fact, if x is any nonnegative eigenvector corresponding to $\rho(A)$, then $x > 0$.

Proof:

Because A is nonnegative, Theorem (I.B.9) guarantees that there is a nonnegative vector x, $x \neq 0$, such that $Ax = \rho(A)x$. If x is any such vector, we have $(I + A)x = [1 + \rho(A)]x$, so $(I + A)^{n-1}x = [1 + \rho(A)]^{n-1}x$. Since A is irreducible, $(I + A)^{n-1} > 0$ by Theorem (I.A.1.13). Hence, $(I + A)^{n-1}x$ is positive by Lemma (I.B.7) since $x \geq 0$ and $x \neq 0$. Therefore, $[1 + \rho(A)]^{n-1}x$ is positive, hence x is a positive vector.

□

(I.B.16) **Lemma**: The geometric multiplicity of an eigenvalue λ of $A \in M_n$ is never greater than its algebraic multiplicity.

Proof: see page 58 in [HJ].

(I.B.17) **Theorem**: If $A \in M_n$ is irreducible and nonnegative, then the eigenvalue $\rho(A)$ has algebraic multiplicity one (hence geometric multiplicity one).

Proof:

We prove by contradiction. If $\rho(A)$ is a multiple eigenvalue of A then $1 + \rho(A)$ is a multiple eigenvalue of $I + A$. But $(I + A)^{n-1} > 0$ hence $\rho[(I + A)^{n-1}] = [1 + \rho(A)]^{n-1}$ has multiplicity one by Theorem (I.B.13).

□

(I.B.18) **Lemma**: Let $A \in M_n$, $A \geq 0$, and suppose A has a positive left eigenvector corresponding to $\rho(A)$. Suppose also that there is another vector x such that $x \geq 0$, $x \neq 0$, and $Ax \geq \rho(A)x$. Then

$Ax = \rho(A)x$.

Proof:

Let y be the given positive left eigenvector of A, i.e. $y^TA = \rho(A)y^T$. Then, $y^T[Ax - \rho(A)x] = \rho(A)y^Tx - \rho(A)y^Tx = 0$. However, $y > 0$ and $Ax \geq \rho(A)x$ by the hypothesis, so it must be the case that $Ax = \rho(A)x$.

□

(I.B.19) **Lemma**: If $A \geq 0$, $x > 0$, and $Ax = 0$ then $A = O$.

Proof:

Each entry of $(Ax)_i$ is zero because $Ax = 0$. In addition, $(Ax)_i = \sum_{k=1}^{n} a_{ik}x_k$ and all $a_{ik} \geq 0$, $x_k > 0$. Therefore, each $a_{ik} = 0$ for $1 \leq i, k \leq n$, or $A = O$.

□

(I.B.20) **Theorem**: Let $A, B \in M_n$. Assume that $A \geq 0$ and irreducible, $|B| \leq A$, $\rho(B) = \rho(A)$, and $\lambda = e^{i\varphi}\rho(B)$ is one of the dominant eigenvalues of B. Then there exist $\theta_1, \ldots, \theta_n \in R$ such that $B = e^{i\varphi}DAD^{-1}$, where $D = \mathrm{diag}(e^{i\theta_1}, \ldots, e^{i\theta_n})$. It follows that the eigenvalues of B are those of A, multiplied by $e^{i\varphi}$.

Proof:

Since $\rho(A) = \rho(B)$, there is some nonzero vector x such that $Bx = \lambda x$ where $|\lambda| = \rho(B) = \rho(A)$. We have $\rho(A)|x| = |\lambda x| = |Bx| \leq |B|\,|x| \leq A|x|$. Further, since A is nonnegative and irreducible, A^T is also. By Theorem (I.B.15), there is a positive vector y such that $A^T y = \rho(A)y$. So, A has a positive left eigenvector, y, corresponding to $\rho(A)$. Apply Lemma (I.B.18) to A with $0 \neq |x| \geq 0$ to see that $A|x| = \rho(A)|x|$. This implies $|x| > 0$ by (I.B.15), and the inequalities above are all qualities. So, $|B|\,|x| = A|x|$, hence $|B| = A$ by Lemma (I.B.19) since $A - |B| \geq 0$.

Since $|\lambda| = \rho(A)$, there is $e^{i\phi}$ such that $\lambda = e^{i\phi}\rho(A)$, Define $\theta_k \in R$ by $e^{i\theta_k} = x_k / |x_k|$ for k = 1, ..., n and $D := \mathrm{diag}(e^{i\theta_1}, \ldots, e^{i\theta_n})$. Then $x = D|x|$ and $BD|x| = Bx = \lambda x = e^{i\phi}\rho(A)D|x|$. Thus,

$e^{-i\phi}D^{-1}BD|x| = \rho(A)|x| = A|x| = |B|\,|x|$. Let $C = e^{-i\phi}D^{-1}BD$ and note that $|C| = |B|$. The preceding equalities say that

$C|x| = |C|\,|x| = |C|x|\,| = \rho(A)|x| > 0$ (because $\rho(A)$ and $|x|$ are both positive). So, for each fixed i, $\left|\sum_{j=1}^{n} c_{ij}|x_j|\right| = \sum_{j=1}^{n} c_{ij}\,|x_j| > 0$, and all $|x_i|$ are positive, hence by Lemma (I.B.11), $|c_{ij}| = c_{ij}$ or $c_{ij} \geq 0$ for all i,j = 1, ..., n, i.e., $C \geq 0$. Hence, $|C| = C$ or $e^{-i\phi}D^{-1}BD = |C| = |B| = A$.

Therefore, $B = e^{i\phi}DAD^{-1}$.

Finally, $D^{-1}BD$ and B have the same eigenvalues with the same multiplicities, and $D^{-1}BD = e^{i\phi}A$, hence the eigenvalues of B are those of A multiplied by $e^{i\phi}$.

□

(I.B.21) **Theorem**: Let $A \in M_n$ be nonnegative and irreducible. If A has

$k \geq 1$ distinct eigenvalues of modulus $\rho(A)$, $\lambda_0, \ldots, \lambda_{k-1}$, then each λ_p has algebraic multiplicity one and they can be ordered so that $\lambda_p = e^{2\pi i p/k}\rho(A)$. Moreover, if A has $m > 1$ eigenvalues of any modulus r less than $\rho(A)$, then k divides m. In fact, the set of eigenvalues of modulus r is invariant under multiplication by $e^{2\pi i/k}$.

Proof:

If $k = 1$, the assertion is trivial. Assume $k > 1$. For $p = 0, \ldots, k-1$, let $0 \leq \varphi_p < 2\pi$ so that $\lambda_p = e^{i\varphi_p}\rho(A)$. Relabel if necessary so that $0 = \varphi_0 < \ldots < \varphi_{k-1} < 2\pi$. Fix p and apply Theorem (I.B.20) by letting $B = A$ and $\lambda = \lambda_p$. Let D_p be the diagonal matrix constructed there so that $A = e^{i\varphi_p} D_p A D_p^{-1}$. Let $\rho = \rho(A)$ and let $\sigma = \{\rho e^{i0}, \rho e^{i\varphi_1}, \ldots, \rho e^{i\varphi_{k-1}}\}$ be the set of dominant eigenvalues of A. By (I.B.20), σ is invariant under multiplication by $e^{i\varphi_p}$, hence σ is invariant under $e^{-i\varphi_p}$ also. Since p was arbitrary, σ is invariant under $e^{i(\varphi_p-\varphi_s)}$ for each p and s between 0 and (k-1). Therefore, σ is invariant under $e^{i\varphi}$, where $\varphi = \min\{|\varphi_p - \varphi_s|, 0 \leq p \neq s \leq k-1\}$. Since ρ is a positive real number, $e^{i\varphi}\rho$ is in σ , and $0 < \varphi \leq \varphi_p$ for all $p = 1, \ldots, k-1$, φ must be equal to φ_1. Also, since $e^{i2\varphi}\rho$ is in σ and $\varphi \leq \varphi_2 - \varphi_1$, 2φ must be equal to φ_2. Continue this argument to see that $(k-1)\varphi = \varphi_{(k-1)}$; therefore φ must be $2\pi//k$.

To complete the proof, suppose $0 < r < \rho$ and let σ' be the set of all eigenvalues of modulus r. By the above, $e^{i\varphi}\sigma' = \sigma'$. If μ is an element of σ', then $\{\mu, \mu e^{i\varphi}, \ldots, \mu e^{i(k-1)\varphi}\}$ is a subset of σ' having k elements, and σ' is the union of disjoint sets of this form, hence the number of eigenvalues in σ' is a multiple of k.

□

(I.B.22) **Remark**:

If A is nonnegative and irreducible and has $k > 1$ eigenvalues of maximum modulus, then all eigenvalues of A are carried to themselves under the rotation $\varphi = 2\pi/k$. Thus, k is a divisor of the number of nonzero eigenvalues of A. In particular, if $0 \leq A \in M_n$ is nonsingular and n is prime, there must be either one or n eigenvalues of maximum modulus. Further, k is a divisor of the number of eigenvalues of any fixed nonzero modulus r.

(I.B.23) **Example**:

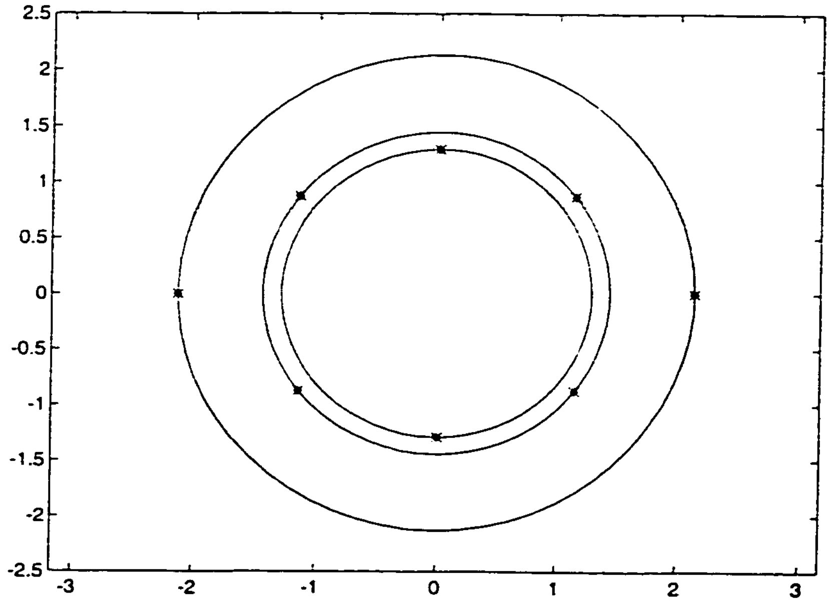
The following 8 x 8 matrix A has two dominant (real) eigenvalues, four subdominant ones of the same modulus not on the real axis, and two more subdominant ones on the imaginary axis (see figure below).

$$A=\left[\begin{array}{cc:cccccc} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ \hdashline 7 & 4 & 0 & 0 & 0 & 0 & 0 & 0 \\ 4 & 7 & 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right]$$ **then $A \geq 0$ and irreducible since $(A + I)^7$ is**

positive:
$$(A+I)^7 = \begin{bmatrix} 3330 & 5768 & 2856 & 1386 & 616 & 224 & 56 & 1640 \\ 224 & 162 & 40 & 32 & 56 & 70 & 56 & 56 \\ 616 & 616 & 162 & 40 & 32 & 56 & 70 & 224 \\ 1368 & 1848 & 616 & 162 & 40 & 32 & 56 & 616 \\ 2586 & 4536 & 1848 & 616 & 162 & 40 & 32 & 1386 \\ 5768 & 9830 & 4536 & 1848 & 616 & 162 & 40 & 2856 \\ 11704 & 20152 & 9830 & 4536 & 1848 & 616 & 162 & 5768 \\ 6952 & 11704 & 5768 & 2856 & 1386 & 616 & 224 & 3330 \end{bmatrix}.$$

The eigenvalues of A are:	**Their moduli are:**
-2.1319	2.1319
2.1319	2.1319
$-1.1506+0.8730i$	1.4443
$-1.1506-0.8730i$	1.4443
$1.1506+0.8730i$	1.4443
$1.1506-0.8730i$	1.4443
$0+1.2917i$	1.2917
$0-1.2917i$	1.2917

They are evenly spaced on the circles centered at (0,0), with radii equal their moduli as below:



Summarizing (I.B.3), (I.B.15), (I.B.17), and (I.B.21) yields the Perron-Frobenius Theorem.

(I.B.24) **Perron-Frobenius Theorem**: If $A \in M_n$ is nonnegative and irreducible, then

(a) $\rho(A) > 0$;

(b) $\rho(A)$ is an eigenvalue of A of algebraic multiplicity one and there is a positive vector **x**, such that $Ax = \rho(A)x$; there is only one such vector having $\|x\|_1 = 1$ and this is called the Perron vector of A.

(c) If A has $k > 1$ distinct eigenvalues of modulus $\rho(A)$, then each has algebraic multiplicity one and they are the k^{th} roots of unity multiplied by $\rho(A)$.

(d) If A has $m > 1$ distinct eigenvalues of modulus r less than $\rho(A)$, then m is a multiple of k; in fact, the set of eigenvalues of modulus r is invariant under multiplication by $e^{2\pi i/k}$.

As we saw in (I.B.13), when A is positive (c) can be replaced by "$\rho(A)$ is the only dominant eigenvalue." This is the classic Perron Theorem (containing part (a), (b), and replaced (c)).

PART TWO

CALCULATING THE PERRON EIGENVECTOR BY UNCOUPLING

INTRODUCTION

A nonnegative matrix $A \in M_n$ is called "nearly uncoupled" if it has a k-level partition for some $k \geq 2$ and the norms of the off diagonal blocks A_{ij} are small compared to the norms of the diagonal blocks A_{ii}. Such matrices appear commonly in many applications. For example, the stochastic matrices which arise in the analysis of queuing networks, computer systems, discrete economic models, and various models from biology and social science often have this property.

In such applications, it is usually important to determine an eigenvector for the dominant eigenvalue. While there are good numerical tools for calculating eigenvalues—e.g., iterative methods—finding accurate values for the corresponding eigenvectors can be very difficult, especially when the matrix is nearly uncoupled. For example, a stochastic matrix A which has a k-level partition that is nearly uncoupled is nearly block-diagonal, and the diagonal blocks are nearly stochastic. So each diagonal block has an eigenvalue very close to one. Thus A has more than one eigenvalue very close to one. Furthermore, a curious fact discovered by Meyer and Hartfiel [MH] is that the second largest eigenvalue of a stochastic matrix can be extremely close to the dominant eigenvalue, even when there is no near uncoupling in the system. We will present here a method developed in [MM] and [M] that shows promise for more accurate calculation of the dominant eigenvector in situations such as these.

A. BACKGROUND

We will give some background about primitivity and M-matrices (section II.A), discuss in detail the [M] paper where Meyer presents his method (section II.B), and will also discuss briefly the [MM] and [HM] papers (section II.B.5).

I. **PRIMITIVITY** [HJ, chapter 8.5]

(II.A.1.1) **Lemma**: Suppose $A \in M_n$ is nonnegative and irreducible and has $k > 1$ eigenvalues of modulus ρ (hence $n > 1$). If m is not an integral multiple of k, then every diagonal entry of A^m is zero. In particular, all $a_{jj} = 0$.

Proof:

If λ is an eigenvalue of A of maximum modulus, then by (I.B.21) $\lambda = e^{i\varphi}\rho(A)$, where $\varphi = 2\pi/k$. Applying Theorem (I.B.20) with $B = A$, we find that $A = e^{i\varphi}DAD^{-1}$ where D is diagonal. Thus, $A^m = e^{im\varphi}DA^mD^{-1}$. So,

$(A^m)_{jj} = e^{im\varphi}(A^m)_{jj}$ for all $1 \le j \le n$. If $(A^m)_{jj} \neq 0$ for any j, then $e^{im\varphi} = 1$, so m is an integral multiple of k. Therefore, $(A^m)_{jj} = 0$ for all j, when m is not an integral multiple of k. In particular, because m = 1 is not an integral multiple of k >1, all $a_{jj} = 0$.

□

(II.A.1.2) **Remark**:

The converse of the above lemma is false. That is, having all diagonal entries zero in a nonnegative irreducible matrix A does not imply there is more than one dominant eigenvalue. For example, if $A = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix}$, then A is irreducible by (I.A.1.13) since $(I + A) > 0$. The diagonal entries of A are zeros but it is easy to check that the eigenvalues of A are 2, −1, −1, so A has only one dominant eigenvalue.

(II.A.1.3) **Definition**: A nonnegative matrix $A \in M_n$ is *primitive* if A is irreducible and has only one eigenvalue of maximum modulus.

(II.A.1.4) **Theorem**: If $A \in M_n$ is nonnegative, then A is primitive if and only if $A^m > 0$ for some $m \ge 1$.

Proof:

Assume $A \geq 0$ and $A^m > 0$. Then as a consequence of Theorem (I.A.1.11), from each node P_i of G(A) there is a path of length m to any other node P_j. Thus G(A) is strongly connected, which is equivalent to A being irreducible (Theorem (I.A.1.16)). Further, A must have only one eigenvalue of maximum modulus. For suppose A has $k > 1$ eigenvalues of maximum modulus; then by Lemma (II.A.1.1), every diagonal entry of A^p is zero when p is not an integral multiple of k. However, $A^m > 0$ for some $m \geq 1$, so for all integers p larger than m, A^p is positive. That is, there is an integer p, not a multiple of k, such that all the diagonal entries of A^p are positive, which contradicts Lemma (II.A.1.1).

Conversely, assume A is primitive; i.e. A is nonnegative irreducible, and has only one eigenvalue of maximum modulus. Then that value is $\rho = \rho(A)$ and it has algebraic, hence geometric, multiplicity one by (I.B.21). Define x, y, $L := xy^T$, and $\rho(A) = \lambda$ as in Theorem (I.B.10). By condition (i) of that theorem, $\left(\frac{1}{\rho}A\right)^m \to L$ as $m \to \infty$. Also, by the Perron-Frobenius Theorem (I.B.24), we can choose $x, y > 0$, hence $L > 0$. Thus, for a sufficiently large integer k, $\left(\frac{1}{\rho}A\right)^k > 0$, hence $A^k > 0$ since $\rho > 0$.

□

(II.A.1.5) **Corollary** [BP, p. 34]: Let $A \in M_n$ be a nonnegative irreducible matrix. If trace(A) is positive then A is primitive.

Proof:

Because A is nonnegative and trace(A) is positive, there is some positive diagonal entry. Let k be the number of maximum modulus eigenvalues of A. If $k > 1$, then by **Lemma** (II.A.1.1), $a_{jj} = 0$ for all $1 \leq j \leq n$. Thus, k must equal one, i.e. A is primitive.

□

(II.A.1.6) **Remark**:

The converse of the above corollary is not true.

Let $A = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix}$. Then $A^2 = \begin{bmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{bmatrix} > 0$, so A is primitive by (II.A.1.4).

However, trace(A) = 0.

II. M – MATRICES

(II.A.2.1) **Definition** [BP, p. 133]: A matrix $A \in M_n$ is an *M-matrix* if

a. $A = [a_{ij}]$, where $a_{ij} \le 0$ for $i \ne j$; and

b. A can be written as $A = sI - B$, where $B \ge 0$, $s > 0$, and $s \ge \rho(B)$.

For example, if $A = \begin{bmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{bmatrix}$, then it can be written as

$A = 2I - \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix} = sI - B$. The matrix B is nonnegative with eigenvalues 2, -1, and -1. Thus, $s \ge \rho(B)$, hence, A is an M-matrix.

Notice s and B are not unique. The matrix A also equals $3I - C$, where

$C = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$; the eigenvalues of C are 3, 0, 0.

(II.A.2.2) **Lemma**: Let $T \in M_n$ be nonnegative. Then T is convergent if and only if $(I - T)^{-1}$ exists and $(I-T)^{-1} = \sum_{m=0}^{\infty} T^m$, where T^0 denotes I. So, if T is nonnegative and convergent, then $(I - T)^{-1}$ is nonnegative.

Proof:

Assume T is convergent. The following argument is based on the

Neumann Expansion [O, p. 201]. Let $S_k = \sum_{m=0}^{k} T^m$. Then $(I - T)\,S_k = I - T^{k+1}$, $k \geq 0$. Since $T^m \to 0$, we see that $(I - T)\,S_k \to I$ as $k \to \infty$. The determinant function and matrix multiplication are continuous, so this implies $\det(I - T)\det(S_k) \to 1$. Therefore, $\det(I - T) \neq 0$, i.e. $(I - T)^{-1}$ exists. Again, using the continuity of matrix multiplication, we see that $(I - T)\,S_k \to I$ implies $(I - T)^{-1}(I - T)\,S_k \to (I - T)^{-1} I$, i.e. $S_k \to (I - T)^{-1}$. That is, $(I - T)^{-1} = \sum_{m=0}^{\infty} T^m$, so $(I - T)^{-1} \geq 0$ is true.

Conversely, assume $(I - T)^{-1}$ exists and $(I - T)^{-1} = \sum_{m=0}^{\infty} T^m \geq 0$. Since T is nonnegative, by the Perron-Frobenius Theorem (I.B.24), $\rho(T)$ is an eigenvalue of T and there is a nonnegative eigenvector x corresponding to $\rho(T)$. Then, $(I - T)x = [1 - \rho(T)]x$, hence $1 - \rho(T) \neq 0$ since $x \neq 0$ and $(I - T)$ is nonsingular. So we have $(I - T)^{-1} x = \frac{1}{1-\rho(T)} x$. Since $(I - T)^{-1} \geq 0$, $x \geq 0$, and $x \neq 0$, we have $\frac{1}{1-\rho(T)} x = (I - T)^{-1} x \geq 0$ and this vector is not zero. Thus, $\frac{1}{1-\rho(T)} > 0$, or $1 > \rho(T)$. Therefore by (I.A.2.7), $T^m \to 0$ as $m \to 0$, i.e. T is convergent.

□

(II.A.2.3) **Theorem**: Suppose $A \in M_n$ is a nonsingular M-matrix of form $A = sI - B$, with $B \geq 0$, $s > 0$, and $s \geq \rho(B)$. Let $\rho = \rho(B)$. Then

a. $s > \rho$ and

b. $A^{-1} = \sum_{m=0}^{\infty} \frac{B^m}{s^{m+1}}$, **hence** $A^{-1} \geq 0$.

Proof:

a. Since B is nonnegative, by Theorem (I.B.9), ρ is an eigenvalue of B, so (s – ρ) is an eigenvalue of A. Then, s ≠ ρ since A is nonsingular, and since s ≥ ρ, we have s > ρ.

b. Because s is positive, $A = sI - B = s(I - s^{-1}B)$. Let $T = s^{-1}B$, so $\rho(T) < 1$ by (a). Then $T^m \to 0$ by Theorem (I.A.2.7) and by Lemma (II.A.2.2) above,

$(I - T)^{-1} = \sum_{m=0}^{\infty} T^m \geq 0$. That is,

$$A^{-1} = \frac{1}{s}\left(I - \frac{1}{s}B\right)^{-1} = \frac{1}{s}(I - T)^{-1} = \frac{1}{s}\sum_{m=0}^{\infty} T^m = \frac{1}{s}\sum_{m=0}^{\infty}\left(\frac{B}{s}\right)^m = \sum_{m=0}^{\infty}\left(\frac{B^m}{s^{m+1}}\right).$$

Hence $A^{-1} \geq 0$ since $B \geq 0$.

□

(II.A.2.4) **Example**:

Let $A = \begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \\ 0 & -1 & 2 \end{bmatrix} = 2I - \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 1 & 0 \end{bmatrix} = sI - B$. Then A is nonsingular and B is nonnegative with eigenvalues 1 and $\frac{1 \pm \sqrt{5}}{2}$. Thus, we see that $s = 2 > \frac{1 + \sqrt{5}}{2} = \rho(B)$ so A is a nonsingular M-matrix. It is easy to see that

$A^{-1} = \begin{bmatrix} 1 & 2 & 1 \\ 0 & 2 & 1 \\ 0 & 1 & 1 \end{bmatrix}$. So $A^{-1} \geq 0$ is true, and it must also be true that $A^{-1} = \sum_{m=0}^{\infty} \frac{B^m}{2^{m+1}}$.

(II.A.2.5) **Definition**: $adj(A) = \left[(-1)^{i+j} A_{ji}\right] = \begin{bmatrix} A_{11} & -A_{21} & \cdots & (-1)^{n+1} A_{n1} \\ -A_{12} & & & \\ \vdots & & \ddots & \\ (-1)^{n+1} A_{1n} & & & A_{nn} \end{bmatrix}$, where

$A_{ij} = \det A(\bar{i}, \bar{j})$, the determinant of the submatrix of A obtained by deleting the i^{th} row and j^{th} column of A.

(II.A.2.6) **Theorem** (Cofactor expansion) [LT p.33]: If $A \in M_n$, then for any fixed i, j ($1 \leq i, j \leq n$), $\det A = \sum_{k=1}^{n} (-1)^{i+k} a_{ik} A_{ik} = \sum_{k=1}^{n} (-1)^{k+j} a_{kj} A_{kj}$.

(II.A.2.7) **Lemma**: Let $A \in M_{m,n}$. Then A has rank r if and only if A has a nonsingular r x r submatrix, and no larger submatrix of A is nonsingular.

Proof:

Suppose that A has rank r. Then, A has r independent columns. Let A_1 be the m x r submatrix whose columns are those r independent columns. So, $\mathrm{rank}(A_1) = r$. That, in turn, has r independent rows, forming a submatrix A_2

(of A_1) of size r x r whose rows are those r independent rows. So, $A_2 \in M_r$ is a submatrix of A and has rank r; therefore, it is nonsingular.

If $r < s \leq m,n$, then any set of s columns of A will be linearly dependent. Thus, if B is any square submatrix of A of size s x s, B will be singular.

Conversely, suppose that A contains a nonsingular r x r submatrix and no larger square submatrix is nonsingular. The first assumption says that A has at least r independent columns; i.e. rank(A) $\geq$ r. The second one means that no more than r columns of A are independent; i.e. rank(A) $\leq$ r. So, rank(A) = r.

□

(II.A.2.8) **Definition** [HJ, p. 16]: For a vector x in C^n, let $\bar{x} = [\bar{x}_i]$, where $\bar{x}_i$ denotes the complex conjugate of x_i. Let $x^* = \bar{x}^T$. Given any subset S of C^n, the *orthogonal complement* of S is the set

$S^\perp := \{ x \in C^n \mid x^*y = 0 \text{ for all } y \in S\}$.

(II.A.2.9) **Lemma**: If $A \in M_n$, then

a. $A[adj(A)] = [adj(A)]A = \det(A)I_n$. So, if A is nonsingular then

$$A^{-1} = \frac{1}{\det(A)} adj(A);$$

b. If A is singular, then either rank(A) < n – 1 and adj(A) = O, or rank(A) = n – 1 and rank(adj(A)) = 1.

Proof:

a. This is well known [L, p. 198].

b. If rank(A) < n − 1, then by Lemma (II.A.2.7), any submatrix of A of size (n−1) x (n−1) is singular. Thus, $\det[A(\bar{i},\bar{j})] = A_{ij} = 0$ for each i, j , hence adj(A) = O.

If rank(A) = n − 1, then some $A_{ij} \neq 0$, i.e. $A_{ij} = \det[A(\bar{i},\bar{j})] \neq 0$. So, rows (1, ..., i−1, i+1, ..., n) of A are independent. Let Row(A) denote the span of the rows of A. Thus, dim(Row(A)) = n − 1, hence $\dim(\mathrm{Row}(A))^{\perp} = 1$. But from (a), $A[\mathrm{adj}(A)] = \det(A)I_n = 0$ because det(A) = 0 since A is singular. Therefore each column of adj(A) is orthogonal to each row of A, which implies each column of adj(A) is in the one-dimensional subspace $(\mathrm{Row}(A))^{\perp}$, hence each column of adj(A) is a multiple of column i in adj(A) (column i is nonzero because $A_{ij} \neq 0$). Thus, rank(adj(A)) = 1.

□

Part (b) of the above lemma can be restated in the following interesting way.

(II.A.2.10) **Corollary**: Let A be an arbitrary n x n matrix. Then,

$$\begin{aligned}\mathrm{rank}(\mathrm{adj}(A)) &= n \text{ when A is nonsingular,}\\ &= 1 \text{ when rank(A)} = n-1,\\ &= 0 \text{ when rank(A)} < n-1.\end{aligned}$$

(II.A.2.11) **Definition**: A *submatrix* of a matrix A is A itself or any matrix that results from deleting some rows and/or columns of A. If the same rows and columns are kept, then the resulting matrix is called a ***principal submatrix*** of A.

For example, let $A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}$. Then some of its submatrices are

$\begin{bmatrix} 1 & 2 & 3 \\ 7 & 8 & 9 \end{bmatrix}$, $\begin{bmatrix} 3 \\ 6 \\ 9 \end{bmatrix}$, and $\begin{bmatrix} 1 & 3 \\ 7 & 9 \end{bmatrix}$, and some of its principal submatrices are

$\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}$, $\begin{bmatrix} 1 & 3 \\ 7 & 9 \end{bmatrix}$ and $[5]$.

(II.A.2.12) **Lemma**: Let $A \in M_n$ be nonnegative and A' be any principal submatrix of A; then $\rho(A') \le \rho(A)$.

In particular, $\max_{i=1,\dots,n} a_{ii} \le \rho(A)$.

Proof:

Let B be the matrix formed from A by putting zeros in the positions that are not in A'. Then B and A' have the same eigenvalues because $A'x' = \lambda x'$ if and only if $Bx = \lambda x$, where x' and x have the same components except $x_j = 0$ when column j is not included in A'. So, $\rho(B) = \rho(A')$. Then $0 \le B \le A$, and $\rho(B) \le \rho(A)$ by (I.B.1), so, $\rho(B) = \rho(A') \le \rho(A)$. When the size of A' is 1 x 1, we have $a_{ii} = \rho(A') \le \rho(A)$. Thus, $\max_{i=1,\dots,n} a_{ii} \le \rho(A)$. □

(II.A.2.13) **Lemma**: Let $A \in M_n$ be a nonsingular M-matrix. Then all principal submatrices are also nonsingular M-matrices.

Proof:

Since A is a nonsingular M-matrix, $A = sI - B$, where $B \geq 0$, $s > 0$, and $s > \rho(B)$, by (II.A.2.3). Let A' be any proper principal submatrix of A and B' be the corresponding submatrix of B (i.e., B' is obtained from B by deleting the same rows and columns as were deleted from A to obtain A'). So $B' \geq 0$ and $\rho(B') \leq \rho(B)$ by (II.A.2.12). Thus, $A' = sI - B'$ and $B' \geq 0$, $s > 0$, and $s > \rho(B) \geq \rho(B')$. So A' has a form of an M-matrix and it is nonsingular because the preceding inequality implies that A' has no zero eigenvalue.

□

(II.A.2.14) **Theorem**: Let $A \in M_n$ be an M-matrix of form $A = sI - B$ where $B \geq 0$, B is irreducible, and $s = \rho(B)$. Then

a. $\mathrm{rank}(A) = n - 1$.

b. $A_{ij} \neq 0$ for all $1 \leq i, j \leq n$.

   In particular, there exist vectors $x > 0$ and $y > 0$ such that $Ax = 0$, $y^T A = 0$ and $\mathrm{adj}(A) = \delta x y^T$ where $\delta \neq 0$. Therefore, either $\mathrm{adj}(A) > 0$ or $\mathrm{adj}(A) < 0$.

c. Any proper principal submatrix is a nonsingular M-matrix.

Proof:

a. B is nonnegative and irreducible, so s is an algebraically and geometrically simple eigenvalue of B by the Perron-Frobenius Theorem (I.B.24). So, $sI - B$ has $(s - s) = 0$ as an algebraically and geometrically simple eigenvalue; therefore, $\mathrm{rank}(sI - B) = n - 1$.

b. By (a) above and Lemma (II.A.2.9), $\mathrm{rank}(\mathrm{adj}(A)) = 1$. Also, by the Perron-Frobenius Theorem applied to B, there is a positive vector x such that $Bx = sx$. So, $Ax = (sI - B)x = sx - sx = 0$. Thus, $x \in [\mathrm{Row}(A)]^{\perp}$.

By the proof of Lemma (II.A.2.9) (b), all columns of adj(A) are in $[\mathrm{Row}(A)]^{\perp}$, which is one-dimensional; thus, each column of adj(A) is a multiple of x.

Since B is nonnegative and irreducible, B^T is also. Further, B^T has the same eigenvalues (with the same multiplicities) as B has. So, there is a positive vector y such that $B^Ty = sy$; thus, $A^Ty = (sI - B^T)y = sy - sy = 0$, or $y^TA = 0$. So, $y^T \in [\mathrm{Col}(A)]^{\perp}$. Using $[\mathrm{adj}(A)]A = O$ from Lemma (II.A.2.9) (a), we see that the rows of adj(A) are in $[\mathrm{Col}(A)]^{\perp}$. Thus, each row of adj(A) is a multiple of y^T.

So, the columns of adj(A) are multiples of x and the rows of adj(A) are multiples of y^T, hence $\mathrm{adj}(A) = \delta xy^T$ for some δ. We also have $\delta \neq 0$ since $\mathrm{rank}(\mathrm{adj}(A)) = 1$, so $A_{ij} \neq 0$ for all i, j. Finally, since $x, y > 0$, $\mathrm{adj}(A) > 0$ if $\delta > 0$ and $\mathrm{adj}(A) < 0$ if $\delta < 0$.

c. Because $A_{ii} \neq 0$ for all $1 \leq i \leq n$, we see that is, all principal submatrices of size (n–1) x (n–1) of A, are nonsingular. As in the proof of (II.A.2.13), all principal submatrices of A are also M-matrices. Any proper principal submatrix of A of a smaller size is a principal submatrix of one of those of size (n–1) x (n–1); hence it is a nonsingular M-matrix by (II.A.2.13).

□

(II.A.2.15) **Example**:

Let A be the matrix in the example following Definition (II.A.2.1), which is a singular M-matrix with $s = \rho(B)$ and B is irreducible. Then all proper principal submatrices of A are nonsingular M-matrices as shown below:

$A'_1 = A'_2 = A'_3 = \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix} = 2I - \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ and $\det(A'_1) = \det(A'_2) = \det(A'_3) = 3$;

$A'_4 = A'_5 = A'_6 = [2] = 2I - [0]$ and $\det(A'_4) = \det(A'_5) = \det(A'_6) = 2$.

B. UNCOUPLING
THE PERRON EIGENVECTOR PROBLEM

I. INTRODUCTION

Let $A \in M_n$ be a nonnegative irreducible matrix with spectral radius ρ. The uncoupling method to determine its unique Perron vector, $v \in C^n$, is this: using a partition of A, we create smaller matrices P_1, P_2, ... P_k of order r_1, r_2, ..., r_k respectively, where $\sum_{i=1}^{k} r_i = n$, so that these smaller matrices have the following properties:

(a) Each smaller matrix P_i is also nonnegative and irreducible, hence each has its unique Perron vector $v^{(i)}$;

(b) Each smaller matrix P_i has the same spectral radius ρ as the original matrix A;

(c) Each $v^{(i)}$ is determined completely independently of the others; and

(d) It is easy to couple the smaller Perron vectors $v^{(i)}$ back together to produce the Perron vector v of the original matrix A.

II. PERRON COMPLEMENTATION

In this section, we introduce the concept of the Perron Complement for a nonnegative irreducible matrix A and develop some of its basic properties. This is based on [M].

(II.B.2.1) **Definition**: Let $A \in M_n$ be nonnegative and irreducible. A *k-level partition of A* is $A = \begin{bmatrix} A_{11} & A_{12} & \cdots & A_{1k} \\ A_{21} & A_{22} & \cdots & A_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ A_{k1} & A_{k2} & \cdots & A_{kk} \end{bmatrix}$ where all diagonal blocks are square. Choose such a partition. Let A_i denote the principal block submatrix of A obtained by deleting the i^{th} row and i^{th} column of blocks. Let A_{i*} denote the i^{th} row of blocks with A_{ii} removed, and A_{*i} denote the i^{th} column of blocks with A_{ii} removed. That is,

$$A_{i*} = [A_{i1} \dots A_{i,i-1}\ A_{i,i+1} \dots A_{ik}] \text{ and } A_{*i} = \begin{bmatrix} A_{1i} \\ \vdots \\ A_{i-1,i} \\ A_{i+1,i} \\ \vdots \\ A_{ki} \end{bmatrix}.$$

The *Perron complement of A_{ii}* in the partitioned A is defined to be the matrix $P_{ii} = A_{ii} + A_{i\bullet}(\rho I - A_i)^{-1}A_{\bullet i}$, where $\rho = \rho(A)$. This term is well defined when $k \geq 2$ and A is nonnegative and irreducible (see (II.B.2.2) (a) below).

For example, the Perron complement of A_{22} in $A = \begin{bmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \\ A_{31} & A_{32} & A_{33} \end{bmatrix}$ is given by $P_{22} = A_{22} + [A_{21}\ A_{23}] \begin{bmatrix} \rho I - A_{11} & -A_{13} \\ -A_{31} & \rho I - A_{33} \end{bmatrix}^{-1} \begin{bmatrix} A_{12} \\ A_{32} \end{bmatrix}$.

(II.B.2.2) **Remark**:

(a) When A has a 2-level partition, the Perron complement has a connection with the classic Schur complement as follows. Consider a 2-level partition of a nonnegative irreducible matrix $A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}$ with spectral radius ρ. The Perron complements of A_{11} and of A_{22} in A are $P_{11} = A_{11} + A_{12}(\rho I - A_{22})^{-1}A_{21}$ and $P_{22} = A_{22} + A_{21}(\rho I - A_{11})^{-1}A_{12}$ respectively. The Schur complement [HJ, p.21] of $(\rho I - A_{11})$ in $(\rho I - A)$ is $S_1 = (\rho I - A_{22}) - A_{21}(\rho I - A_{11})^{-1}A_{12}$. So, $P_{22} = \rho I - S_1$. Similarly, if S_2 denotes the Schur complement of $(\rho I - A_{22})$ in $(\rho I - A)$, then $P_{11} = \rho I - S_2$.

(b) When A is nonnegative, the matrix $(\rho I - A)$ is a singular M-matrix, but if A is also irreducible, all its proper principal submatrices are nonsingular M-matrices, by (II.A.2.14). Therefore, for a nonnegative and irrducible A, when $k \geq 2$, the inverse of $(\rho I - A_i)$ exists and $(\rho I - A_i)^{-1}$ is nonnegative for each $i = 1, \ldots, k$.

Assume in all that follows that A is nonnegative and irreducible.

(II.B.2.3) **Lemma**: Suppose A has a k-level partition and let $\widetilde{A} = QAQ$ where Q is the permutation matrix which corresponds to an interchange of the 1st and ith block positions. (For such Q, $Q^{-1} = Q^T = Q$.) The matrix $\widetilde{A}$ can be repartitioned into a 2 x 2 block matrix $\widetilde{A} = \begin{bmatrix} \widetilde{A}_{11} & \widetilde{A}_{12} \\ \widetilde{A}_{21} & \widetilde{A}_{22} \end{bmatrix}$ so that $\widetilde{A}_{11} = A_{ii}$ and the Perron complement of A_{ii} in A is the same as the Perron complement of $\widetilde{A}_{11}$ in $\widetilde{A}$. That is, $P_{ii} = \widetilde{P}_{11} = \widetilde{A}_{11} + \widetilde{A}_{12}(\rho I - \widetilde{A}_{22})^{-1}\widetilde{A}_{21}$.

Proof:

Partition the n x n identity matrix as $\begin{bmatrix} I_1 & & O \\ & \ddots & \\ O & & I_k \end{bmatrix}$ so that for each p, I_p

has the same shape as A_{pp}. Let Q = $\begin{bmatrix} O & & \cdots & & I_i & & \cdots & O \\ & I_2 & & & & & & \\ \vdots & & \ddots & & \vdots & & & \vdots \\ & & & I_{i-1} & & & & \\ I_1 & & \cdots & & O & & \cdots & \\ & & & & & I_{i+1} & & \\ \vdots & & & & & & \ddots & \vdots \\ O & & \cdots & & & & & I_k \end{bmatrix}$.

Then QAQ has the form described.

□

(II.B.2.4) **Theorem**: Let A $\in$ M_n be nonnegative and irreducible with spectral radius ρ and consider a k-level partition of A. Then

a. Each Perron complement P_{ii} is also a nonnegative irreducible matrix whose spectral radius is again ρ.

b. If the Perron vector v of A is partitioned conformably as

$v = \begin{bmatrix} v^{(1)} \\ v^{(2)} \\ \vdots \\ v^{(k)} \end{bmatrix}$, then each $v^{(i)}$ is a positive eigenvector for P_{ii}

associated with ρ, i.e. $P_{ii}\, v^{(i)} = \rho v^{(i)}$.

c. Write e for a column vector of ones, of whatever length is appropriate where it appears. Then $p_i = \dfrac{v^{(i)}}{e^T v^{(i)}}$ is the Perron vector of P_{ii} for each i.

Proof:

a. By (II.B.2.3), $P_{ii} = \widetilde{P}_{11} = \widetilde{A}_{11} + \widetilde{A}_{12}(\rho I - \widetilde{A}_{22})^{-1}\widetilde{A}_{21}$, and by (II.B.2.2)(b), $(\rho I - \widetilde{A}_{22})^{-1} \geq 0$. So, $\widetilde{A}_{12}(\rho I - \widetilde{A}_{22})^{-1}\widetilde{A}_{21}$ is nonnegative, hence P_{ii} is nonnegative since it is a sum of nonnegative matrices. We will show P_{ii} is irreducible by showing its graph $G(P_{ii})$ is strongly connected, which suffices by Theorem (I.A.1.14).

Suppose first that i = 1, A_{11} is r x r, and A has been partitioned to the form $A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}$, so $P_{11} = A_{11} + A_{12}(\rho I - A_{22})^{-1}A_{21}$. To see that the graph $G(P_{11})$ is strongly connected, let E(.) denotes the set of directed paths in the directed graph of a specified matrix and let h, k $\in$ {1, ..., r}. We will show that there is a directed path from node h to node k in $G(P_{11})$. Notice that P_{11} is the sum of the *nonnegative* matrices A_{11} and $A_{12}(\rho I - A_{22})^{-1}A_{21}$, so $E(A_{11}) \subseteq E(P_{11})$ and $E(A_{12}(\rho I - A_{22})^{-1}A_{21}) \subseteq E(P_{11})$.

Case 1: There is a path from node h to node k in $G(A_{11})$. Then there is a path from node h to node k in $G(P_{11})$; hence, P_{11} is strongly connected.

Case 2: There is no path from node h to node k in $G(A_{11})$. However, there is a path from node h to node k in G(A) since A is irreducible. This path must pass through at least one node i $\in$ {r+1, ..., n}. Suppose the path leaves

$G(A_{11})$ just once. That is the path in $G(A_{11})$ is of the type

$$\underbrace{hh_1, \ldots, h_{p-1}h_p}_{in\, E(A_{11})}, h_p i_1, \ldots, i_q k_1, \underbrace{k_1 k_2, \ldots, k_s k}_{in\, E(A_{11})} \text{ and}$$

$\varnothing \neq \{i_1, i_2, \ldots, i_q\} \subseteq \{r+1, \ldots, n\}$. The existence of the directed edges $h_p i_1, \ldots, i_q k_1$ in $E(A)$ guarantees that $[A_{12}(A_{22})^{q-1} A_{21}]_{h_p k_1} \neq 0$ by (I.A.1.9).

Further, $(\rho I - A_{22})$ is a nonsingular M-matrix so $(\rho I - A_{22})^{-1} = \sum_{j=0}^{\infty} \frac{(A_{22})^j}{\rho^{j+1}} \geq 0$ by (II.A.2.3). Thus, $A_{12}(\rho I - A_{22})^{-1} A_{21} = \sum_{j=0}^{\infty} \frac{A_{12}(A_{22})^j A_{21}}{\rho^{j+1}}$; hence

$[A_{12}(A_{22})^{q-1} A_{21}]_{h_p k_1} \neq 0$ implies $[A_{12}(\rho I - A_{22})^{-1} A_{21}]_{h_p k_1} \neq 0$,

so that $(h_p i_1, \ldots, i_q k_1)$ is a path in $G(A_{12}(\rho I - A_{22})^{-1} A_{21})$, so in $G(P_{11})$ as well.

Thus, the path $\underbrace{hh_1, \ldots, h_{p-1}h_p}_{in\, E(A_{11})}, \underbrace{h_p i_1, \ldots\ldots\ldots, i_q k_1}_{in\, E(A_{12}(\rho I - A_{22})^{-1} A_{21})}, \underbrace{k_1 k_2, \ldots, k_s k}_{in\, E(A_{11})}$ will be in $G(P_{11})$.

Now suppose the path from h to k leaves $G(A_{11})$ more than once. Then it will be the concatenation of several paths, each of the type above. Each of those is in $G(P_{11})$, hence the longer path is also.

Suppose now that $i \neq 1$. By Lemma (II.B.2.3), there is a permutation matrix Q so that $Q^{-1} = Q$ and $\widetilde{A} = QAQ = \begin{bmatrix} \widetilde{A}_{11} & \widetilde{A}_{12} \\ \widetilde{A}_{21} & \widetilde{A}_{22} \end{bmatrix}$ where $\widetilde{A}_{11} = A_{ii}$ and $\widetilde{P}_{11} = P_{ii}$. Apply the above to $\widetilde{A}$ and conclude that P_{ii} is irreducible.

Below is a proof that $\rho(P_{11}) = \rho$, which will complete the proof of (a).

b. Define $\tilde{v}$ = Qv and write $\tilde{v} = Qv = \begin{bmatrix} \tilde{v}^{(1)} \\ \tilde{v}^{(2)} \end{bmatrix}$ where $\tilde{v}^{(1)} = v^{(i)}$. Since Av = ρv

and $Q^2 = I$, we can conclude that

$$\rho\tilde{v} = \rho Qv = QAv = QAQ^2v = \tilde{A}\,\tilde{v}\text{, hence } \begin{bmatrix} \tilde{A}_{11}\tilde{v}^{(1)} + \tilde{A}_{12}\tilde{v}^{(2)} \\ \\ \tilde{A}_{21}\tilde{v}^{(1)} + \tilde{A}_{22}\tilde{v}^{(2)} \end{bmatrix} = \begin{bmatrix} \rho\tilde{v}^{(1)} \\ \rho\tilde{v}^{(2)} \end{bmatrix}.$$

Setting the second blocks equal implies $\tilde{v}^{(2)} = (\rho I - \tilde{A}_{22})^{-1}\tilde{A}_{21}\tilde{v}^{(1)}$ and

therefore,

$$P_{ii}v^{(i)} = \tilde{P}_{11}\tilde{v}^{(1)} = \tilde{A}_{11}\tilde{v}^{(1)} + \tilde{A}_{12}(\rho I - \tilde{A}_{22})^{-1}\tilde{A}_{21}\tilde{v}^{(1)} = \tilde{A}_{11}\tilde{v}^{(1)} + \tilde{A}_{12}\tilde{v}^{(2)} = \rho\,\tilde{v}^{(1)} = \rho v^{(i)}.$$

The next-to-last equality follows from the equality of the first blocks in the vector equation above.

Since $\tilde{v} > 0$, this proves that ρ is an eigenvalue of P_{ii} with an associated positive eigenvector $v^{(i)}$. Then by (I.B.6), ρ is indeed the spectral radius of P_{ii}.

c. It follows from (b).

□

III. UNCOUPLING AND COUPLING THE PERRON VECTOR

(II.B.3.1) **Definition**: Let A be a nonnegative irreducible matrix, partitioned at k-level, for some $k \geq 2$, and let v be the Perron vector of A. For each i, let $v^{(i)}$ be defined as in Theorem (II.B.2.4) and let $p_i = \frac{v^{(i)}}{e^T v^{(i)}}$, so p_i is the Perron vector of the Perron complement P_{ii}. The normalizing scalar $\alpha_i = e^T v^{(i)} > 0$ will be called the i^{th} *coupling factor*.

(II.B.3.2) **Observation**:

The Perron vector v of the original matrix A can be written in terms of the Perron vectors p_i of P_{ii} as $v = \begin{bmatrix} \alpha_1 p_1 \\ \alpha_2 p_2 \\ \vdots \\ \alpha_k p_k \end{bmatrix}$.

So, in theory, it is possible to find the Perron vector of a nonnegative irreducible matrix $A \in M_n$ using Perron complementation. That is, if the matrix A is partitioned to k levels and P_{ii} is defined for each diagonal block, the Perron vectors p_i can be determined independently for each P_{ii}, so, if we knew the coupling factors, we would have the Perron vector of A.

However, the coupling factors have the form $\alpha_i = e^T v^{(i)}$, so, it appears that we need to know the vector v first. Fortunately, it turns out that is not the case. In this section, we will show how the coupling factors α_i's can be determined without prior knowledge of $v^{(i)}$'s, hence the technique of uncoupling-coupling does work to yield the Perron vector of A.

(II.B.3.3) **Definition**: Let $A \in M_n$ be nonnegative and irreducible with spectral radius ρ. Assume that A is partitioned to k levels, $k \geq 2$, let P_{ii} be the Perron complement of A_{ii} as in Definition (II.B.2.1), and let p_i denote the Perron vector for P_{ii}. Then the *coupling matrix* associated with the k-level partition of A is $C \in M_k$ whose entries are given by $c_{ij} = e^T A_{ij} p_j$.

We will now show that this matrix C, which is smaller than A, can be used to calculate the coupling factors.

(II.B.3.4) **Theorem**: Let $A \in M_n$ be a nonnegative and irreducible matrix with spectral radius ρ, partitioned to k levels, and let

$$v = \begin{bmatrix} v^{(1)} \\ v^{(2)} \\ \vdots \\ v^{(k)} \end{bmatrix}$$ denote its Perron vector, partitioned conformably. Then

the associated k x k coupling matrix C is also a nonnegative irreducible matrix with the same spectral radius ρ. Let $\alpha = \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_k \end{bmatrix}$

where its components α_i are the coupling factors defined earlier as $\alpha_i = e^T v^{(i)}$. Then $C\alpha = \rho\alpha$, i.e., α is the Perron vector of C. The vector α is called the *coupling vector associated with the partitioned matrix A*.

Proof:

Let p_i denote the Perron vector of P_{ii}, for each i. By definition, $C = [c_{ij}] = e^T A_{ij} p_j$, so C is nonnegative. Now, note that $e^T > 0$, $A_{ij} \geq 0$, and $p_j > 0$. So, $c_{ij} = 0$ if and only if $A_{ij} = 0$. Thus if C could be permuted to a block triangular form, so could A. This is a contradiction since A is irreducible. Thus, C must be irreducible also.

For i = 1, ..., k, $[C\alpha]_i = \sum_{j=1}^{k} c_{ij}\alpha_j = \sum_{j=1}^{k} e^T A_{ij} p_j \alpha_j = e^T \sum_{j=1}^{k} A_{ij} p_j \alpha_j$. Since

$p_j = \dfrac{v^{(j)}}{\alpha_j}$ and $Av = \rho v$, we have $p_j\alpha_j = v^{(j)}$ and hence

$[C\alpha]_i = e^T \sum_{j=1}^{k} A_{ij} p_j \alpha_j = e^T \sum_{j=1}^{k} A_{ij} v^{(j)} = e^T \rho\ v^{(i)} = \rho\alpha_i$. So, $C\alpha = \rho\alpha$. Since C is nonnegative and irreducible, α is positive, and $C\alpha = \rho\alpha$, the spectral radius of

C is ρ, by (I.B.6). Since v is normalized, $\sum_{j=1}^{k} \alpha_j = \sum_{j=1}^{k} e^T v^{(j)} = e^T v = 1$, so the

coupling vector α = $\begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_k \end{bmatrix}$ is normalized and hence is the Perron vector of C.

□

The theorem above together with the observation (II.B.3.2), which says

that $v = \begin{bmatrix} \alpha_1 p_1 \\ \alpha_2 p_2 \\ \vdots \\ \alpha_k p_k \end{bmatrix}$, yields the following theorem, which is the major conclusion

of this paper.

(II.B.3.5) **Theorem** (The coupling theorem): Let A be a nonnegative

irreducible matrix with spectral radius ρ, and let

A = $\begin{bmatrix} A_{11} & A_{12} & \cdots & A_{1k} \\ A_{21} & A_{22} & \cdots & A_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ A_{k1} & A_{k2} & \cdots & A_{kk} \end{bmatrix}$ be a k-level partition with square diagonal

blocks, k ≥ 2. If p_i is the Perron vector of the Perron complement

$P_{ii} = A_{ii} + A_{i*}(\rho I - A_i)^{-1}A_{*i}$ and if $\alpha = \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_k \end{bmatrix}$ is the Perron vector for

the k x k coupling matrix C defined by $c_{ij} = e^T A_{ij} p_j$, then $v = \begin{bmatrix} \alpha_1 p_1 \\ \alpha_2 p_2 \\ \vdots \\ \alpha_k p_k \end{bmatrix}$

is the Perron vector of A. Conversely, if the Perron vector of A is

given in the form of a conformably partitioned vector $v = \begin{bmatrix} v^{(1)} \\ v^{(2)} \\ \vdots \\ v^{(k)} \end{bmatrix}$,

then $\alpha_i = e^T v^{(i)}$ and $\dfrac{v^{(i)}}{\alpha_i}$ is the Perron vector for the Perron

complement P_{ii}.

The following corollary shows that it is especially easy to uncouple and couple the Perron vector when one has a 2-level partition.

(II.B.3.6) **Corollary**: Let A be a nonnegative irreducible matrix with

spectral radius ρ. If A is partitioned as $A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}$ where A_{11}

and A_{22} are square, then the Perron vector of A is given by

$v = \begin{bmatrix} \alpha_1 p_1 \\ \alpha_2 p_2 \end{bmatrix}$, where p_1 and p_2 are the Perron vectors of the Perron complements $P_{11} = A_{11} + A_{12}(\rho I - A_{22})^{-1}A_{21}$ and $P_{22} = A_{22} + A_{21}(\rho I - A_{11})^{-1}A_{12}$, respectively, and where the coupling factors are given by $\alpha_1 = \dfrac{e^T A_{12} p_2}{\rho - e^T A_{11} p_1 + e^T A_{12} p_2}$ and $\alpha_2 = 1 - \alpha_1$.

Proof:

By multiplying out the matrix product

$$C\alpha = \begin{bmatrix} e^T A_{11} p_1 & e^T A_{12} p_2 \\ e^T A_{21} p_1 & e^T A_{22} p_2 \end{bmatrix} \begin{bmatrix} \dfrac{e^T A_{12} p_2}{\rho - e^T A_{11} p_1 + e^T A_{12} p_2} \\ 1 - \dfrac{e^T A_{12} p_2}{\rho - e^T A_{11} p_1 + e^T A_{12} p_2} \end{bmatrix}$$

, one can verify that $C\alpha = \rho\alpha$, i.e. $\alpha = \begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix}$ is the Perron vector of C. Then by Theorem (II.B.3.5) above, we are done.

□

(II.B.3.7) **Remark**:

The preceding simple corollary is the basis for an efficient parallel processing procedure for calculating the Perron vector of A, using a 2-level partition at each stage (roughly partitioning each successive matrix in half at each step). We call this method *divide-and-conquer*. This procedure also minimizes the maximum size of matrix inverses embedded in the Perron complements, which is important for efficient calculation of the coupling factors.

The benefits of this method are further investigated in [MM] and [MH], which we will present in section (II.B.5).

(II.B.3.8) **Example**:

The following example illustrates this divide-and-conquer procedure, using a 2-level partition at each stage. All the computations were done in Matlab.

Example:

Let $A = \left[\begin{array}{cccc|cccc} 0 & 0 & 0 & 1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 & 0 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ \hline 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \end{array}\right]$ be partitioned into $A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}$. Then

the spectral radius of A is s = 2. 0755 and the Perron complements of A are

$$P_{11} = A_{11} + A_{12}(\rho I - A_{22})^{-1}A_{21} = \left[\begin{array}{cc|cc} 0 & 0.6275 & 0.4818 & 1.6275 \\ 0 & 0.6275 & 1.4818 & 0.6275 \\ \hline 1 & 1.1457 & 0 & 0.1457 \\ 0 & 0 & 0.4818 & 0 \end{array}\right] \text{ and}$$

$$P_{22} = A_{22} + A_{21}(\rho I - A_{11})^{-1}A_{12} = \left[\begin{array}{cc|cc} 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \\ \hline 1 & 0.3023 & 0.7732 & 1.3252 \\ 0 & 0.6275 & 0.6047 & 0.7503 \end{array}\right].$$

Now apply the same procedure to the diagonals blocks of P_{11}, obtaining

$[P_{11}]_{11} = \begin{bmatrix} 0.421 & 1.1099 \\ 0.7971 & 1.5407 \end{bmatrix}$, and $[P_{11}]_{22} = \begin{bmatrix} 1.714 & 1.5573 \\ 0.4818 & 0 \end{bmatrix}$. The associated

Perron vectors and the coupling factors of these two later matrices are

$[v_{11}]_1 = \begin{bmatrix} 0.4015 \\ 0.5985 \end{bmatrix}$, $[v_{11}]_2 = \begin{bmatrix} 0.8116 \\ 0.1884 \end{bmatrix}$, $[\alpha_{11}]_1 = 0.6038$, and $[\alpha_{11}]_2 = 0.3962$.

Thus, the Perron vector of P_{11} is

$$v_1 = \begin{bmatrix} 0.6038\begin{bmatrix} 0.4015 \\ 0.5985 \end{bmatrix} \\ 0.3962\begin{bmatrix} 0.8116 \\ 0.1884 \end{bmatrix} \end{bmatrix} = \begin{bmatrix} 0.2424 \\ 0.3614 \\ 0.3215 \\ 0.0746 \end{bmatrix}.$$

Similarly, we obtain

$[P_{22}]_{11} = \begin{bmatrix} 1.4333 & 1.3327 \\ 1 & 0 \end{bmatrix}$, $[P_{22}]_{22} = \begin{bmatrix} 1.3252 & 1.3252 \\ 0.7503 & 0.7503 \end{bmatrix}$,

$[v_{22}]_1 = \begin{bmatrix} 0.6748 \\ 0.3252 \end{bmatrix}$, $[v_{22}]_2 = \begin{bmatrix} 0.6385 \\ 0.3615 \end{bmatrix}$, $[\alpha_{22}]_1 = 0.3131$, and $[\alpha_{22}]_2 = 0.6869$.

The Perron vector of P_{22} is $v_2 = \begin{bmatrix} 0.3131\begin{bmatrix}0.6748\\0.3252\end{bmatrix} \\ \\ 0.6869\begin{bmatrix}0.6385\\0.3615\end{bmatrix}\end{bmatrix} = \begin{bmatrix}0.2113\\0.1018\\0.4386\\0.2483\end{bmatrix}$.

Using the formula in (II.B.3.6) again for computing the coupling factors α_i of A, we have the coupling vector for A is $\alpha = \begin{bmatrix}0.6158\\0.3842\end{bmatrix}$. Thus, the Perron vector of A is

$$v = \begin{bmatrix} 0.6158\begin{bmatrix}0.2424\\0.3614\\0.3215\\0.0746\end{bmatrix} \\ \\ 0.3842\begin{bmatrix}0.2113\\0.1018\\0.4386\\0.2483\end{bmatrix}\end{bmatrix} = \begin{bmatrix}0.1493\\0.2225\\0.198\\0.046\\0.0812\\0.0391\\0.1685\\0.0954\end{bmatrix}.$$

When "format long" is used, the above vector v looks like $\begin{bmatrix}0.14929399519912\\0.22254886064441\\0.19800544110173\\0.04596622299812\\0.08118052938592\\0.03911404459401\\0.16848879781631\\0.09540210826037\end{bmatrix}$.

To check these results, we typed eig(P_{11}) and eig(P_{22}), finding that the dominant eigenvalue for each is 2.0755, as it should be; and the command

[V D] = eig(A) produced the Perron vector identical to the vector v shown above, when displayed using “format long.”

IV. PRIMITIVITY ISSUE

When the dominant eigenvalue of any matrix A is unique, this usually makes numerical calculation of an associated eigenvector easier and more accurate. The classic "power" method gives insight about why. This method is: choose arbitrary $v^{(1)}$ and then calculate $v^{(2)} = Av^{(1)}$, $v^{(3)} = Av^{(2)}$, etc. (Actually, each $v^{(i)}$ is usually scaled somehow before being used to calculate $v^{(i+1)}$.) For most matrices, this iteration will converge to an eigenvector corresponding to the dominant eigenvalue of A. However, writing $v^{(1)}$ as linear combination of eigenvectors reveals why this method can converge slowly, or not at all if the dominant eigenvalue is multiple or if there are other eigenvalues with nearly the same magnitude [W, pp. 570-572].

Therefore the Perron complement method for nonnegative irreducible A is likely to work well if the P_{ii} matrices are primitive. The following theorems show that this can be true even when A itself is not primitive.

(II.B.4.1) **Theorem**: Let A be a nonnegative irreducible matrix partitioned at k-level, $k \geq 2$. If a particular diagonal block A_{ii} is primitive, then the corresponding Perron complement P_{ii} must also be primitive.

Proof:

Since A_{ii} is nonnegative and $A_{i*}(\rho I - A_i)^{-1}A_{*i}$ is nonnegative (see proof of (II.B.2.4)), it follows that for each positive integer n,

$(P_{ii})^n = [A_{ii} + A_{i*}(\rho I - A_i)^{-1}A_{*i}]^n = (A_{ii})^n + N$ where N is a nonnegative matrix. Therefore, $(P_{ii})^n$ is a positive matrix whenever $(A_{ii})^n$ is, i.e. P_{ii} is primitive whenever A_{ii} is (II.A.1.5).

□

(II.B.4.2) **Theorem**: Let A be a nonnegative irreducible matrix with a k-level partition, $k \geq 2$. If a particular diagonal block A_{ii} has at least one nonzero diagonal entry, then the corresponding Perron complement P_{ii} must be primitive.

Proof:

If A_{ii} has at least one nonzero diagonal entry, then so does P_{ii}, hence trace(P_{ii}) is positive Thus, P_{ii} is primitive by Corollary (II.A.1.5), since P_{ii} is nonnegative irreducible with positive trace.

□

(II.B.4.3) **Remark**:

(a) The converse of the above theorem is false. For example, let

$$A = \left[\begin{array}{ccc|ccc} 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 \\ \hline 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 \end{array}\right].$$ Then, A is irreducible (since $(I + A)^4$ is positive),

but none of A_{11}, A_{22}, and A itself is primitive (the eigenvalues of A are -2.1736, 2.1736, $-0.3797 \pm 0.7117i$, and $0.3797 \pm 0.7117i$). Nevertheless, the corresponding Perron complements $P_{11} = \begin{bmatrix} 0.4601 & 0.4601 & 0.9201 \\ 0.9201 & 0.4601 & 0.9201 \\ 0.4601 & 0.9201 & 0.9201 \end{bmatrix}$ and

$P_{22} = \begin{bmatrix} 0.4601 & 0.9201 & 0.4601 \\ 0.4601 & 0.4601 & 0.9201 \\ 0.4601 & 0.9201 & 0.9201 \end{bmatrix}$ are each primitive (by the classic Perron Theorem or by Theorem (II.A.1.4)).

(b) The primitivity of a matrix A does not guarantee that all its Perron complements are primitive. For example, let $A = \left[\begin{array}{cc|c} 0 & 1 & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} \\ \hline 1 & 0 & 0 \end{array}\right]$. The matrix A is irreducible since $(I + A)^2$ is positive and primitive since A^5 is positive. However, for the indicated partition, the Perron complement $P_{11} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ is not primitive.

V. APPLYING THE PERRON COMPLEMENT METHOD TO STOCHASTIC MATRICES

Here we will discuss the results in [MM] and [MH] where the divide-and-conquer method, implemented on parallel processors, is compared to classic methods for calculating the Perron vector of a stochastic matrix.

(II.B.5.1) **Definition** [L, p. 282]: A nonnegative column vector whose entries sum to one is called a *probability vector*. A matrix $A \in M_n$ is called a *stochastic matrix* if its columns are probability vectors. Such A has $\rho(A) = 1$. A *Markov chain* is a sequence of probability vectors $v^{(0)}, v^{(1)}, v^{(2)}, \ldots$, together with a stochastic matrix A such that $Av^{(0)} = v^{(1)}, Av^{(1)} = v^{(2)}, Av^{(2)} = v^{(3)}, \ldots$. An *irreducible Markov chain* is a Markov chain whose stochastic matrix A is irreducible.

For an irreducible Markov chain, the Perron vector of A is called a "stationary distribution vector" because $Av = v$. We will see here that this vector can be found by Perron complementation in a particularly efficient way: when A is stochastic, the associated coupling matrix C is stochastic also, and this makes the divide-and-conquer method defined in (II.B.3.7) work especially well.

(II.B.5.2) **Theorem**: Let $A \in M_n$ be nonnegative, irreducible, and stochastic.

Then

a. The coupling matrix C as defined in (II.B.3.3) is stochastic.

b. For a 2-level partition of A, the coupling vector is

$$\alpha = \frac{1}{c_{12} + c_{21}} \begin{bmatrix} c_{12} \\ c_{21} \end{bmatrix}.$$

Proof:

a. For each j= 1, ..., k, the sum of entries of the j^{th} column of C is given by

$\sum_{i=1}^{k} c_{ij} = \sum_{i=1}^{k} e^T A_{ij} p_j = \left(\sum_{i=1}^{k} e^T A_{ij} \right) p_j = e^T p_j = 1$. Thus, C is stochastic by Definition (II.B.5.1).

b. Assume A has a 2-level partition. Let C be the coupling matrix and let α be the vector defined above. By multiplying out the matrix product

$C\alpha = \begin{bmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{bmatrix} \left(\frac{1}{c_{12} + c_{21}} \right) \begin{bmatrix} c_{12} \\ c_{21} \end{bmatrix}$, one can verify that $C\alpha = \alpha$, i.e. α is the Perron vector of C, which we call the coupling vector.

□

In the paper [MM], the authors discuss operation counts (i.e. flop count), CPU time, speedup (the ratio of CPU time required by an algorithm to execute on a single processor and the time required to execute the same problem on more than one processor), and efficiency (speedup divided by the number of processors) for the divide and conquer method versus other methods like LU and QR. (The power method is not included in their study since it has a low rate of convergence.) Their conclusion is that even though the divide-and-conquer method is not as fast as LU or QR, the result can be more accurate in cases where LU and QR do not work so well.

Call a matrix "nearly uncoupled" if A has a k-level partition for some $k \geq 2$ and the norms of the off diagonal blocks A_{ij} are small compared to the norms of the diagonal blocks A_{ii}. Such matrices appear commonly in many applications such as: discrete economic models, various models from biology and social science, and analysis of queuing networks and computer systems [MH].

If a stochastic matrix A is nearly uncoupled, A is nearly block-diagonal, and the diagonal blocks are nearly stochastic. So each diagonal block has an eigenvalue very close to one (the dominant eigenvalue). This can also happen when A has no near uncoupling. For example, if

$$A = \begin{bmatrix} 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0.0000000001 & 0.9999999999 & 0 & 0 & 0 \\ 0.9999999999 & 0.0000000001 & 0 & 0 & 0 \end{bmatrix}$$, then A possesses an eigenvalue

whose modulus is close the dominant eigenvalue. (The eigenvalues of A are:

1, –0.9999999995, –0.49999999998333 ± 0.86602540375557i, and 0.99999999991667.)

As already observed, any time a matrix has two eigenvalues close to being dominant, the power method is slow to converge. Other methods, such as ones based on LU or QR factorizations can lead to major numerical inaccuracies.

In [MH] the authors examine still another method: perturbing each diagonal block to make it stochastic, calculating the Perron vectors of these smaller matrices, and using coupling to estimate the Perron vector of A. They show this can also give inaccurate results. These discoveries all suggest that the divide-and-conquer method based on Perron complementation can be valuable for stochastic matrices, even though it is more costly to implement than classic methods.

SUMMARY

For an n x n nonnegative irreducible matrix A, the Perron-Frobenius Theorem describes its properties as

(a) $\rho(A)$ is a positive dominant eigenvalue with algebraic (hence geometric) multiplicity one;

(b) There is a corresponding positive eigenvector which is unique when normalized, and this is called the Perron vector of A;

(c) If A has $k \geq 1$ distinct eigenvalues of modulus $\rho(A)$, then each has algebraic multiplicity one and they are the k^{th} roots of unity multiplied by $\rho(A)$;

(d) If A has $m \geq 1$ distinct eigenvalues of any modulus r less than $\rho(A)$, then m is a multiple of k; in fact, the set of eigenvalues of modulus r is invariant under multiplication by $e^{2\pi i/k}$.

When a nonnegative and irreducible matrix, whose spectral radius is ρ, is partitioned to k-level with square diagonal blocks as

$$A = \begin{bmatrix} A_{11} & A_{12} & \cdots & A_{1k} \\ A_{21} & A_{22} & \cdots & A_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ A_{k1} & A_{k2} & \cdots & A_{kk} \end{bmatrix}, \text{ and } v = \begin{bmatrix} v^{(1)} \\ v^{(2)} \\ \vdots \\ v^{(k)} \end{bmatrix}$$ is the conformably partitioned

Perron vector of A, the Perron complement of A_{ii} is defined as

$P_{ii} = A_{ii} + A_{i*}(\rho I - A_i)^{-1}A_{*i}$, $i = 1, \ldots, k$. These are shown to have the following properties:

(a) Each Perron complement P_{ii} is nonnegative and irreducible and its spectral radius is again ρ.

(b) The Perron vector p_i of P_{ii} is the normalized i^{th} segment of the Perron vector v of A. That is, $p_i = \dfrac{v^{(i)}}{\alpha_i}$, where the normalizing factors (called the coupling factors) are $\alpha_i := e^T v^{(i)}$.

c. The k x k coupling matrix C defined as $C = [c_{ij}] = e^T A_{ij} p_j$ is also a nonnegative irreducible matrix with spectral radius ρ and its Perron vector is $\alpha = \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_k \end{bmatrix}$.

d. The above three results allow the Perron vector p_i for each P_{ii} to be determined independently of each other. Then by using the coupling vector α, these p_i can be coupled together to form the Perron vector v of A, by $v = \begin{bmatrix} \alpha_1 p_1 \\ \alpha_2 p_2 \\ \vdots \\ \alpha_k p_k \end{bmatrix}$.

REFERENCES

BP A. Berman and R. J. Plemmons, *Nonnegative Matrices in The Mathematical Sciences*, SIAM, Philadelphia, 1994.

HJ R. A. Horn and C. R. Johnson, *Matrix Analysis*, Cambridge University Press, New York, 1985.

L D. C. Lay, *Linear Algebra and Its Applications*, Addison-Wesley, Menlo Park, California,. 2nd ed., 1997.

LT P. Lancaster and M. Tismenetsky, *The Theory of Matrices*, Academic Press Inc., Orlando, Florida, 2nd ed., 1985.

M C. D. Meyer, Uncoupling the Perron eigenvector problem, *Linear Algebra and Its Applications* **114/115** (1989), 69-94.

MH C. Meyer and D. Hartfiel, On the structure of stochastic matrices with a subdominant eigenvalue near 1, *Linear Algebra and Its Applications* **272** (1998), 193-203.

MM C. Meyer and R. Mattingly, *Numerical Solution of Markov Chains*, ed. W. J. Stewart, Marcel Dekker, Inc., New York, 1991.

O J. M. Ortega, *Matrix Theory*, Plenum Press, New York, 1989

W J. Wilkinson, *The Algebraic Eigenvalue Problem*, Clarendon Press, Oxford, 1965.